STRUCTURAL STUDIES OF TWO PROTEINS

by

Lindsay Sawyer

A thesis presented for the degree of Doctor
of Philosophy of the University of Edinburgh.

Department of Natural Philosophy
University of Edinburgh

June, 1971

Huge trunks! and each particular trunk a growth
Of intertwisted fibres serpentine
Up-coiling, and inveterately convolved.

Wordsworth: Yew Trees.

ABSTRACT

This work can be divided into two parts: in the first, the crystallisation and preliminary X-ray data for rabbit muscle aldolase (E.C. 4.1.2.13) is examined, in the second, the search for suitable high resolution derivatives for bovine β-lactoglobulin is studied with a view to the correlation of solution and X-ray studies of the binding of gold, platinum and iridium complexes to the protein.

Rabbit muscle aldolase from the back and leg muscles of freshly killed rabbits can be prepared by ammonium sulphate fractionation as a microcrystalline suspension. At least two forms of crystal exist: at pH values below 7.0 the crystals are hexagonal bipyramids, above 7.0 hexagonal plates can be formed. Crystals of a suitable size for X-ray work were grown from 2.0M phosphate, pH 6.0 in the form of hexagonal bipyramids. Precession photographs along and perpendicular to the sixfold axis allowed the cell dimensions and space group to be determined as a=163.5 Å, c=335.0 Å and $P6_122$. There were 18 molecules in the unit cell or 1.5 in the asymmetric unit implying that the molecule is a tetramer rather than a trimer. Although crystals of the high pH form are obtained they are not of sufficient quality to allow X-ray study.

With β-lactoglobulin, the binding of tetracyanoaurate (III) is examined in solution by ultraviolet spectroscopy and polarimetry and also by X-rays in the solid phase. In solution it is found that one mole of heavy atom complex per subunit binds specifically and reversibly, Protein - SH + $Au(CN)_4^- \rightleftharpoons$ Protein - S - $Au(CN)_3^-$ + HCN, to the free cysteine causing a marked change in specific rotation, to a stable conformation midway between the R and S states. The effect of concentrated salt is to inhibit this change and it is

proposed that lattice X crystals correspond to the N state, lattices Y and Z to an intermediate one and that the R state should be obtained by crystallising the β-lactoglobulin-gold complex. X-ray studies on lattice Y show that the binding occurs mainly at the free sulphydryl group and also at the $HgI_4^=$ site but in lattice Z, no sulphydryl binding is detected. The derivatives are markedly non-isomorphous.

With the isostructural tetracyanoplatinite (II), no binding to the free sulphydryl group is detected either in solution or by X-ray crystallography. The electron density map shows several minor sites.

Further studies with hexacyanoplatinate (IV) and the chloro-complexes of platinum and iridium show that the iridium complexes look promising as high resolution derivatives although no satisfactory correlation is obtained between solution and crystal studies. The chlorocomplexes of platinum do show evidence of binding in solution but they appear to react to ratios greater than one to one. No satisfactory X-ray data are obtained in lattice Z with these derivatives.

In conclusion, a strategy for the crystallisation of proteins with a view to the requirements of the crystallographer is proposed.

TABLE OF CONTENTS

CHAPTER 6

ABBREVIATIONS

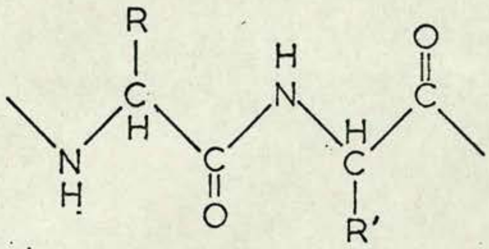
NOTE: This list is not complete. It contains only those abbreviations used which are not defined when they are first used.

DAP, DHAP	Dihydroxyacetone phosphate
DTNB	5,5'-Dithiobis(2-nitrobenzoic acid)
FDP	Fructose-1,6, Diphosphate
GAP, G3P	Glyceraldehyde-3-phosphate
GDH	Glycerol-1-phosphate dehydrogenase
GDP, GLP	Glycerol-1-phosphate
G-3-PDH	Glyceraldehyde-3-phosphate dehydrogenase
NAD^+, NADH	Oxidised and reduced forms of nicotinamide adenine dinucleotide
pCMB	p-chloromercuribenzoate
pCMBS	p-chloromercuribenzoate sulphonic acid
TIM	Triose phosphate isomerase

CHAPTER I

A. General Introduction

Proteins are a remarkable class of compounds: they can be structural like keratin or ~~chitin~~ collagen or they can be functional, in the chemical sense, acting as catalysts, regulators and protectors. But they are all polymers made up of twenty α-amino acids linked in the same way by the peptide bond.



It is the variation in the side chains, R, R' and so on, which confers the unique properties on each protein. The particular sequence of residues in the polypeptide chain is under genetic control, the protein being the eventual translation of deoxyribonucleic acid, according to the genetic code, into a product capable of carrying out the cell's specific need at the time. The whole process is complex and not fully understood but it is likely that the protein produced has some means of controlling its own production so that a balance is maintained within the cell. One line of research directed at aiding comprehension of these processes is that of studying protein structure.

With the burgeoning of high resolution protein structures in the last five or six years, the time is fast approaching when some general principles regarding the structure and function of proteins (and more particularly, enzymes) from the molecular viewpoint, will be able to be put forward. At the present time when perhaps a dozen structures have been determined by X-ray crystal structure analysis (see reviews by Blow and Steitz (1970), North and Phillips (1969) and Dickerson and Geis (1968)), there is still not really enough information on the various facets of protein structure to form any such principles with certainty. However, Blow and Steitz (1970) do go so far as to say that two facts seem to have emerged from the structures already determined. These are (i) in most cases charged groups are on the surface of the molecule, uncharged ones in the interstices, unless a charged one is associated with the activity of the protein and (ii) almost all hydrogen-bond donors are conveniently placed for hydrogen-bond acceptors.

Most structure determining effort in recent years has been put into the study of enzymes, the catalysts responsible for the control of metabolism an accurate knowledge of whose specificity and mode of action is vital if life-processes are to be understood fully. But what sort of information can be obtained from a study of enzymes on a molecular level?

The way in which substrate bonds are made and cleaved by any particular enzyme can be more easily determined if the structure of the active site is known in detail. Since an efficient method of catalysis is heterogeneous and because most cell processes take place in aqueous solution, the enzyme has to contain a heterogeneous centre for the catalysis (where any concentration of charge is not 'diluted' by the water) and at the same time remain soluble in water. This is done by the nature of the twenty amino-acids which enable a polypeptide chain of these units to fold in such a way as to push the hydrophobic residues to the centre whilst leaving the polar ones on the outside in such a way as to maintain solubility. Consequently, the primary sequence is very important in that it determines the overall shape, although the actual method of folding as the chain is produced from the ribosome is not yet understood. Also, it is this relationship between the primary sequence and the tertiary and even quaternary structures about which little is known. It <u>is</u> known, however, that the replacement of only one amino-acid residue in one region can completely alter the characteristics of the protein whilst in another, no effect is noticed (e.g. Perutz and Lehman (1968)). With the enzyme, lysozyme, Phillips and his coworkers have been able to show the binding of the substrate to the active site, situated in a cleft on the molecule (Phillips (1966b)). From this they have been able to complement the solution studies on the mechanism of action.

Moving from one enzyme to a group of them, Shotton, Watson and Hartley (1970) have shown that the primary sequences of porcine elastase and bovine α-chymotrypsin A and B are very similar and that many of the differences are conservative (i.e. hydrophobic residue for hydrophobic residue and so on). Further, a comparison of the structures of elastase and α-chymotrypsin shows that they are broadly similar, any small change in the internal structure at one point being offset by another small change close by. These small changes do, however, alter the specificities of the enzymes. But caution must be exercised in the drawing of parallels between similarly-sequenced proteins without having direct structural evidence. A tentative model for the structure of α-lactalbumin, part of the lactose synthetase system in milk, was proposed on the basis of the similarity between the sequences of lysozyme and the milk protein (Browne et al. (1969)). Of the 125 or so residues 45 are identical, 23 are conservative and the disulphide bridges (4) are in the correct places, but although the structure seems appropriate, no activity towards the substrate in the absence of the other part, galactosyl transferase, is exhibited. Is the active site of α-lactalbumin distorted into the active configuration by association with the other part?

Another point of great interest is raised by the structure of subtilisin (Kraut et al. (1969)). This is a bacterial enzyme which attacks the same substrates as α-chymotrypsin. The sequences of the active region are shown below, the active serine residue in both enzymes being underlined.

<u>α-chymotrypsin</u>: -Cys-Gln-Gly-Asp-<u>Ser</u>-Gly-Gly-Pro-Val-Val-Cys-
195

<u>Subtilisin</u>: -Tyr-Asn-Gly-Thr-<u>Ser</u>-Met-Ala-Ser-Pro-His-Val
221

One main difference shown by the above short lengths of chain is that in the former sequence there are two disulphide bridges close to the residue responsible for the activity whilst in the latter there are no such links in that part of the chain. A comparison of the overall three-dimensional structures shows that there is little in common between them yet they act on similar substrates. Apparently, therefore, evolution has found two possible answers to the same problem. Is this sort of situation more common than one might expect and, conversely, might it be possible to "tailor" an enzyme to meet a specific need, in a sense bypassing evolution?

In living organisms, certain metabolic pathways exist for the synthesis of essential substances and for the degradation of, amongst other things, food. These pathways take a substance and, by a series of steps, convert it into something quite different. Are the enzymes responsible for each stage related in some simple way to those performing adjacent steps? Since enzymes, as catalysts, are capable of performing both forward and reverse reactions, two adjacent ones in a pathway must have similar substrate binding sites for at least part of the molecule. For example, in the glycolytic pathway for the breakdown of starch to lactic acid, the enzymes aldolase and glyceraldehyde-3-phosphate dehydrogenase have an active lysine residue responsible for binding part of the substrate. What will the structures of the enzymes be like; will there be any similarity between them?

Looking further ahead, Blow and Steitz (1970) ask the following two questions. Certain polar groups within a protein are known to have abnormal pK_a values; β-lactoglobulin, for instance, has a carboxyl group which is titrated in the same region as imidazole residues (Nozaki, Bunville and Tanford (1959)). Might it be possible to predict such changes from the primary structure which is responsible for imposing a certain environment on a particular residue? Will the structures and detailed mechanisms even be sufficient to explain everything about proteins and enzymes, or will there arise other problems which have been unimagined so far?

The answers to these and the many other questions raised must wait at least until the structures of more proteins have been determined, which, it is hoped will be in the not too distant future.

The last few paragraphs have discussed in very general terms why the structural study of proteins is important; the next question which can be asked is how can the structures best be examined? In the last thirty years, many sophisticated advances have been made in the study of proteins, both in solution and in the solid state. Although techniques such as polarimetry and ultraviolet (u/v) spectroscopy were known before this, their application to proteins had been limited. Nowadays, still the only way of obtaining complete tertiary structural data is by X-ray crystallography but this does not mean that valuable information about the structure cannot be determined by other methods. Since the basis of this work is crystallographic in outlook, a fuller discussion of the X-ray methods used for proteins is given in the second part of this introduction. The remainder of this section will therefore be devoted to a brief discussion of the information about proteins of direct use to the crystallographer, which can be obtained from other techniques.

Perhaps the most useful of all studies which can supplement that of crystallography, is the determination of the primary structure. Most of the specific techniques involved in this process are detailed in Methods in Enzymology (Vol. 11, 1967) and, more recently, by Schroeder (1968). Basically the method involves the hydrolysis of the polypeptide chain, either chemically or enzymatically, the separation of the fragments and then the careful removal and identification, one at a time, of the individual residues from these more manageable fragments. The use to which the sequence can be put by the crystallographer is in the interpretation of the electron density map. Since some of the side-chains are of a similar size and shape, for example valine and threonine, as far as their distinction by X-ray techniques is concerned, interpretation of the protein Fourier map is very tricky indeed. If, however, the sequence is known and certain key features can be correlated (like the presence of a cysteine close to the binding site of a Hg atom, a disulphide bridge or the indole group of a tryptophan) then the sequence can be "fitted" to the map. Recently, the advent of mass spectrometers capable of resolving high molecular weight molecules and the computer analysis of the spectra obtained from such instruments has enabled peptides to be sequenced (Sheldrick (1970)).

Binding studies of various kinds can be of great use also, in the preparation of heavy atom* derivatives for the X-ray analysis. With enzymes, substrates and substrate analogues can be used to great effect for this labelling process. For instance, in α-chymotrypsin toluene sulphonic acid and p-iodo-benzene sulphonic acid were used as a "native" and derivative pair which label the reactive serine (Matthews et al., (1967)). Other examples of

---

* Throughout this thesis the phrases "heavy atom" and "heavy atom derivative" should be taken as implying "a complex containing a heavy metal atom" and "a protein to which a heavy atom complex has been attached", respectively. The phrases are used because at low resolution only the heavy metal atom nucleus of the complex shows up in the difference maps.

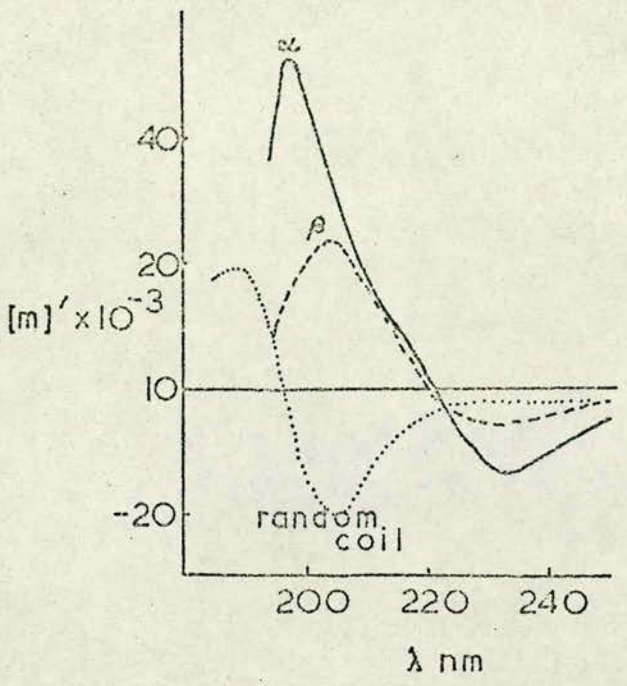


FIG. I.1 The ORD spectra of poly-L-lysine in water, Sarkar and Doty (1966).

binding include the use of mercurials for attachment to sulpydryls (e.g. Dunnhill and Green (1965)).

Moving from the more chemical methods to the more physical ones, infrared (i/r), u/v and visible spectroscopy and optical rotatory dispersion (ORD) and circular dichroism (CD) can be used, the latter pair mainly to obtain information on the conformation or secondary structure of the protein. Pauling and his coworkers (1951) proposed three types of basic secondary structure for proteins: the α-helix, the pleated sheet and the random coil structures. Any solution methods for determining the relative amounts of these three types of structure require "calibration" with poly-amino-acids and known tertiary structures. This has been done to some extent, but the results are still little better than qualitative. The X-ray evidence points to there being rather more distorted β-structure than was hitherto realised (Blow and Steitz (1970)). A most useful method for examining solution conformations is ORD. Fig. I. 1 shows the data of Sarkar and Doty (1966) for poly-L-lysine in its three forms where m' is given by

$$[m'] = (3/n^2+2).\ (MRW/100).[\alpha]_\lambda = a_o\lambda_o^2/(\lambda^2-\lambda_o^2) + b_o\lambda_o^4/(\lambda^2-\lambda_o^2)^2$$

where n is the refractive index of the solution and MRW is the mean residue molecular weight. The second equation is that of Moffitt and Yang (1957) and has been found to describe the shapes of the ORD curves obtained experimentally. a_o is a measure of the residue rotations and the interactions within the helix and is dependent on the solvent. b_o and λ_o are independent of the solvent and are functions of the helical backbone of the polypeptide chain. The figure shows the differences in the shape of the curves arising from the different conformations of the polymer in solution. This data was confirmed independently by Davidson, Tooney and Fasman (1966). The following year, Greenfield, Davidson and Fasman (1967) showed that by computing curves for various ratios of the three types of structure and comparing these with the observed data it was possible to predict the amounts of α, β and random structure in proteins and polypeptides. Magar (1968) has given a more elegant method for doing this by least squares rather than by trial and error. Still, however, the method suffers from the existence in proteins of aromatic groups, disulphide bridges and distorted forms of α- and β-structure which are not adequately described by the equation used for the fitting of theory to experiment. When the effects of these groups are known then a much better fit should be obtained.

Circular dichroism, the difference in absorption of left and right circularly polarised light, has recently been used to great effect in "unscrambling" the overlapping Cotton curves obtained by ORD in the u/v regions. This is because it is an absorptive effect and hence rapidly falls to zero away from an absorption band. Consequently, ORD and CD can be used in a complementary fashion. Greenfield and Fasman (1969) have used computed CD curves as a means of determining the relative proportions of the various types of secondary structure. Although they claim the best fit of experiment and theory to be in cases where there is a high degree of non-random secondary structure as in myoglobin, the results for the more irregular proteins are encouraging and, in fact, the use of CD is a decided improvement over ORD in the evaluation of protein conformation.

I/r spectroscopy can also be used to give information about the conformation of proteins. The effect of the conformation of the poly-peptide chain on the various amide bands has been reviewed recently by Hardy (1969) and, more fully, by Jencks (1963). This method can be used to distinguish between parallel and anti-parallel pleated sheet.

Finally, the aggregation of protein subunits into oligomers, that is the quaternary structure, can be examined by various non-crystallographic methods which can help the three dimensional structure determination. Hydrodynamic methods like ultracentrifugation, electrophoresis, and chromatography can give values for the molecular weight, the overall dimensions and the number and molecular weights of the subunits. This information can be of great assistance in the checking of the crystallographic determination of molecular weight and space group. The electron microscope has been used to show directly the subunit association of proteins with molecular weights greater than about 100,000 (e.g. Penhoet et al. (1967)). A review by Klotz, Langerman and Darnall (1970) discusses the quaternary structures of proteins more fully.

Briefly, then, these are the main ways in which supplementary information can be obtained for a structural study of a protein. A problem is posed, however, by the very method used for the determination of the tertiary structure, which does not affect the other studies in the same way. How justified is the extrapolation from the static, time-average structure obtained by X-ray crystallography to that of the protein in vitro, if not actually in vivo? Some justification derives from the recent paper by Rossi and Bernhard (1970) on the intramolecular hydrolysis of an α-chymotrypsin-substrate complex. They find that the rates of hydrolysis at various pH values are the same, within experimental error, in solution and in the solid state. Moreover, the electronic spectra associated with the enzyme-substrate complex in both cases are almost the same showing that the environments of the chromophore (which is also the substrate) are almost identical at an electronic level. They conclude that the structure in the crystal is a very close approximation to that in solution.

B. Structure Determination of Proteins by X-ray Diffraction

The following account is only meant as a broad outline, fuller accounts being available in reviews by Phillips (1966), Holmes and Blow (1966) and North and Phillips (1969).

The reconstruction of a molecular structure from its diffraction pattern can be regarded as comprising two stages: the production of an electron density map and then the interpretation of this to give an image of the structure. The latter is aided by a knowledge of the amino-acid sequence, partial or complete, but it can be done with reasonable accuracy without such information (Lipscomb et al. (1966)). The former stage is performed using the expression for the electron density:

$$p(xyz) = \frac{1}{V}\sum_h\sum_k\sum_l \underline{F}(hkl).\exp(-2\pi i(hx+ky+lz)) \quad (1)$$

where x,y,z are the fractions of the unit cell edges at which the electron density is calculated, V is the cell volume, h,k,l are the Miller indices of the reflections used in the summation and $\underline{F}(hkl)$ is the structure factor which can be expressed as:

$$\underline{F}(hkl) = \sum_{j=1}^{N} f(hkl)_j.\exp(2\pi i(hx_j+ky_j+lz_j)) \quad (2)$$

where $f(hkl)_j$ is the atomic scattering factor of the jth. atom whose fractional coordinates are x_j,y_j,z_j, there being N atoms in the unit cell. $\underline{F}(hkl)$ is considered as being made up of $|F(hkl)|$, the amplitude of the wave scattered from the hkl planes, which can be obtained from the observed intensity, and $\alpha(hkl)$, the phase of this wave relative to that from ~~(000)~~ the origin, which unfortunately cannot be measured and must be calculated. If the positions of all of the atoms in the asymmetric unit, and hence the unit cell, are known (i.e. x_j,y_j,z_j with j=1,2,--,N in equation (2)) then this value can be calculated. However, the atomic coordinates are what is required so that there is an immediate problem: how can the phases be obtained? This is the main part of the calculation of the electron density map and for proteins, it is done by the method of isomorphous replacement.

The Isomorphous Replacement Method.

The isomorphous replacement method as a means of phase determination was suggested by Bokhoven, Schoone and Bijvoet (1951) who pointed out that the phases could be determined from

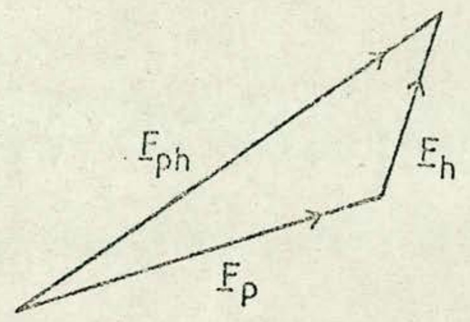


FIG. I.2 Vector triangle, $\underline{F}_{ph} = \underline{F}_p + \underline{F}_h$

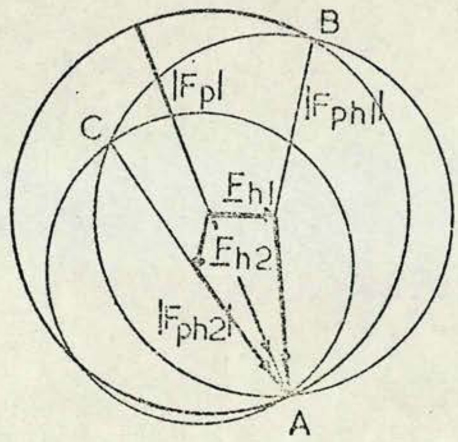


FIG. I.3 Idealised solution of the phase problem with two derivatives.

a series of three or more isomorphous crystals with differing heavy atom structures. It was not until the paper by Green, Ingram and Perutz (1954) that it was realised that the method was applicable to proteins and that it could, in fact, yield phase information. Up to the present time, it has been responsible for the solution of the dozen or so protein structures so far determined and now is being used as the only sure method of obtaining high resolution electron density maps.

In essence, the method is as follows. In Fig. I. 2, the general case of the protein structure factor, $\underline{F}_P$, and those of the derivative (i.e. protein plus heavy atom), $\underline{F}_{PH}$, and heavy atom contribution alone, $\underline{F}_H$, are depicted as a vector triangle. This corresponds to the equation

$$\underline{F}_{PH} = \underline{F}_P + \underline{F}_H \tag{3}$$

However, only $\underline{F}_H$ is fully known (see below), the magnitudes, $|F_{PH}|$ and $|F_P|$ being all that is known of the other two vectors. This allows circles to be drawn of radius $|F_P|$ and $|F_{PH}|$ at either end of the vector $\underline{F}_{H1}$, as shown in Fig. I. 3, the points of intersection, A and B, being the two possible solutions of the vector triangle. This ambiguity can be resolved if another heavy atom derivative is prepared having a different heavy atom site, represented by $\underline{F}_{H2}$. A similar construction gives intersections at A and C thus affording a unique solution, A, to the problem.

This is a highly idealised representation of the problem and, in practice, it is advisable to have as many derivatives as possible since, owing to the various errors which arise from intensity measurement and lack of isomorphism, the 'phase triangle' seldom closes exactly and as many indications as possible of the correct phase are therefore desirable.

The method, then, can be divided into five separate stages:

(a) preparation of an isomorphous series of crystals (i.e. a native and several derivatives),

(b) the collection of complete data sets for each of these,

(c) the determination of the positions of the heavy atoms, their reference to the same origin and the refinement of their parameters,

(d) the use of the heavy atom vectors for the calculation of the phases

(These four stages will be dealt with in more detail in the subsequent sections) and

(e) the use of these phases to calculate the electron density map by means of one the standard programs available for this operation.

The Preparation of Derivatives.

The problems associated with the crystallisation of the native protein will be dealt with in Chapter III. The preparation of an isomorphous series of crystals now appears to be the limiting stage in the three-dimensional structure determination of proteins. Once a native crystal form with neither too large a unit cell nor too many molecules per asymmetric unit, has been obtained, the search for suitable derivatives can begin.

There is, as yet, no certain method available for producing derivatives although, broadly, three possible approaches can be tried. These are described by Holmes and Blow (1966) and are the following:

(a) chemical modification followed by crystallisation,

(b) crystallisation in the presence of the reagent

and (c) diffusion of the reagent into the crystal after it has grown. Method (a) might be of use where specific reactions can occur but very often, previous attachment of a ligand can either cause the protein to crystallise in a different form or even stop it crystallising altogether. (b) has been found to be poor for, if reaction between protein and heavy atom complex can occur, then the objections are as for (a) and, if crystals are obtained, they are most likely better prepared by method (c). Method (c) has been the most successful method of preparation. Because a protein is typically 50% mother liquor, it is possible to diffuse into the lattice fairly large heavy atom molecules which can then bind to the protein. Unfortunately, no really systematic way of doing this has been found and much still depends on trial and error although there are usually several compounds, e.g. pCMBS and K_2HgI_4, which are tried first because of previous success with them. A fuller account of the preparation of derivatives appears in a recent review by Blake (1968).

As a rough measure of the effect of a particular additive, Crick and Magdoff (1956) have suggested

$$\langle (\Delta I)^2 \rangle^{\frac{1}{2}} / \langle I \rangle \doteqdot \sqrt{2}.(N_e/N_p)^{\frac{1}{2}}.(f_e/f_p)$$

where I is the native intensity, ΔI is the change caused by the addition, N_e is the number of added atoms per unit cell, N_p is the number of protein atoms in the cell and f_e, f_p are the scattering factors of the heavy and mean protein atoms respectively.

However, it is based on the use of Wilson statistics which assume a random distribution of scattering material throughout the unit cell, an assumption which is inaccurate for proteins (Harker (1953), Luzzati (1955)).

Lack of isomorphism can be detected most easily by a change in cell dimensions but more subtle changes can take place which are not reflected in marked changes in the 'external' isomorphism. Such changes, caused by alteration of the protein molecules' relative arrangement on binding the heavy atom complex, can often be detected (Carlisle and Palmer (1962)) by examining the difference in the intensity distributions as a function of sinΘ as suggested by Crick and Magdoff. More sophisticated tests have been devised by Parthsarathy and Ramachandran (1966) but these have not yet been used in practice. Indeed, Phillips (1966a) doubts if they would be sensitive enough for proteins.

Data Collection

Data collection from protein crystals requires two problems to be considered which are not normally encountered with small molecules. The first is that the crystal must be kept moist during exposure,which requires sealing it in a glass capillary by the method of Boyes-Watson, Davidson and Perutz (1947) or King (1954). This raises the second problem by making absorption correction awkward, but a satisfactory method has been devised by North, Phillips and Matthews (1968).

The Scaling of Protein and Derivative Data.

Before proceeding with the determination of the heavy atom positions, it is necessary to put the various sets of data on the same relative scale as the native. It is only possible at this stage to put the native data on an absolute scale by comparison with a standard crystal but, in this method, provided the various sets of data are on the same relative scale, the structure can be determined.

The usual method employed is to put

$$k.\sum_{h\,k\,l} |F_{PH}|^2 = \sum_{h\,k\,l} |F_P|^2$$

where k $\doteqdot$ 1.05 and is readily refined at a later stage. This has been found quite adequate although, in cases with noticeable fall-off at high values of sinΘ, a scale factor of the form

$$k = K_\circ \exp(-D.\sin^2\theta/\lambda^2)$$

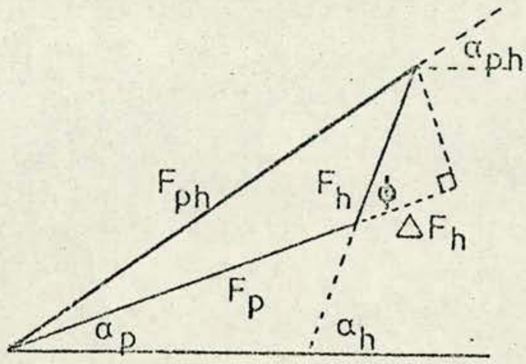
can be used, where K and D are determined from a plot of $\ln(\overline{|F_P|}/\overline{|F_{PH}|})$ versus $\sin^2\theta$ (see, for example, Dickerson et al. (1967)).

Determination of Heavy Atom Positions.

The method used to locate the heavy atoms is one similar to the difference Patterson synthesis. Theoretically, this method is more likely to give interpretable results if a centrosymmetric projection is available because in this case the phases are either 0 or π. Thus, the usual vector equation, (3), can be written as

$$\underline{F}_H = \pm |F_{PH}| \mp |F_P| \qquad (4)$$

Usually, $|F_{PH}|$ and $|F_P|$ will have the same sign (i.e. the heavy atom contribution is relatively small) so that a synthesis with coefficients $(F_{PH}-F_P)^2$, rather than $|F_{PH}|^2 - |F_P|^2$ which has undesirable cross-vectors (those between native protein and heavy atom), should contain only the heavy atom-heavy atom vectors



<u>FIG. I.4</u> Vector triangle showing the approximation $\Delta F_h = F_h . \cos \phi$.

allowing the sites to be located (Green, Ingram and Perutz (1954)). In the case which can arise when both $|F_{PH}|$ and $|F_P|$ are small and have opposite sign, a "crossover", the coefficient should really be $(F_{PH}+F_P)^2$ but because of the awkwardness of preparing semi-substituted derivatives to detect such crossovers and because their number is usually small, the error introduced by neglecting them does not interfere significantly with the main features.

For non-centrosymmetric data, however, this expression, (4), is only an approximation. Fig. I. 4 shows that there is no need for α_{PH} and α_P to be the same. However, Blow (1958) has shown that a synthesis using $(|F_{PH}| - |F_P|)^2$ for the non-centric [100] projection of horse haemoglobin does yield the heavy atom vectors, although the peaks appear with half weight and the background is higher but diffuse (Ramachandran (1964)). From Fig. I. 4, it can be seen that $\emptyset = \alpha_H - \alpha_P$ and, <u>if</u> $\alpha_{PH}-\alpha_P$ is small then $\emptyset$ is small also and

$$|\Delta F_H| = |F_{PH}| - |F_P| \doteqdot |F_H|.\cos\emptyset$$

Thus,

$$|\Delta F_H|^2 = |F_H|^2.\cos^2\emptyset$$

$$\doteqdot \tfrac{1}{2}.|F_H|^2 + \tfrac{1}{2}.|F_H|^2.\cos.2\emptyset$$

The first term gives the heavy atom peak of half weight, the second the diffuse background. Expression (4) is the standard one used for finding the heavy atoms both in two- and three-dimensional syntheses. However, Rossman (1961) has shown that a Patterson summation with coefficients $(|F_{PH}(hkl)| - |F_{PH}(\overline{hkl})|)^2$ will reveal the positions of the anomalous scatterers, assuming that there are only a few such atoms in the unit cell. This led Kartha and Parthasarathy (1965) to show that an inclusion of the anomalous dispersion data should lower the background still assuming that $\alpha_{PH} - \alpha_P$ is small. A similar treatment to the above gives

$$|\Delta F_{ano}| \doteq |F_H|.\sin\phi$$

so that

$$|\Delta F_{iso}|^2 + |\Delta F_{ano}|^2 \doteq |F_H|^2$$

Matthews (1966b) has taken this a stage further by allowing the value of $\alpha_{PH} - \alpha_P$ to take any value up to 90^o. That is

$$|F_P| + |F_{PH}| \gg |F_H|$$

a condition apt not to be obeyed by weak reflections. Singh and Ramaseshan (1966) have removed all assumptions and have given an exact solution which is rather cumbersome, in terms of $|F_P|^2$ and $|F_m|^2$ where $|F_m|^2 = \frac{1}{2}(|F^+_{PH}|^2 - |F^-_{PH}|^2)$. In practice, however, it is usually possible to locate the major heavy atom sites by use of the $|\Delta F|^2$ synthesis. In fact, it has been found that a three-dimensional vector map is easier to interpret than a projection because of the latter's likelihood of containing overlapping peaks (Green and Komorowski, unpublished).

Correlation Functions.

Before attempting to calculate phases with the heavy atom sites obtained by the method above, it is necessary to ensure that their coordinates are all referred to the same origin. This is more likely to occur in the lower symmetry space groups where the origin is not necessarily fixed by the symmetry. For example, in haemoglobin crystallised from ammonium sulphate (A.S.) solution, the crystals are monoclinic, space group C2.

Blow (1958) was faced with the problem that, whilst the x and z coordinates of each of the three derivatives, p-chloromercuribenzoate, dimercuriacetic acid and mercuric acetate were known and from the [100] projections the y coordinates could be found, there was no way of relating them to the same origin. That is, the choice of origin along the y-axis is arbitrary. However, Blow was able to make use of two functions worked out by Perutz (1956) in order to determine the relative positions along y. Although these did work, they are seldom used because of more recent developments. Rossman (1960) showed that a synthesis calculated with coefficients

$$(|F_{PH1}| - |F_{PH2}|)^2$$

gives a direct indication of the relationship between the two derivatives in that the heavy atom (1) - heavy atom (2) vector appears as a negative peak. The background is of the same general type as is found in the $(\Delta F)^2$ synthesis except that the heavy atom self-vectors are present as positive peaks as well. An improvement on this is to use the function suggested by Steinrauf (1963)

$$(|F_{PH1}| - |F_P|)(|F_{PH2}| - |F_P|)$$

This gives only the heavy atom (1)-heavy atom (2) vector as a positive peak.

Refinement of the Heavy Atom Sites.

Also, before calculating the phases it is necessary to obtain the best values for the heavy atom parameters as is possible. The most common method of refinement for centrosymmetric data is that of trial and error as developed by Hart (1961). This minimises the value of

$$E = \sum((\pm k.\ |F_{PH}^{obs}| \pm |F_P^{obs}|) - \underline{F}_H^{calc})^2$$

by applying fixed shifts, one at a time, to the positional parameters, occupancy, scale factor and temperature factor. As this method was employed extensively throughout this work, a fuller account is given in Appendix I. A full least-squares modification of this method has been described by Lundberg (1965).

For non-centric data, the refinement really requires phase information. Kraut et al. (1962) proposed a method where the positional parameters, occupancies and scale factor for each derivative were refined separately using the current estimate of the phase calculated from all of the derivatives. Each cycle of refinement was followed by a recalculation of the phases. The function being minimised was

$$\sum_{hkl} (k.|F_{PH}^{obs}| - |F_{PH}^{calc}|)^2 = \sum_{hkl} (k.|F_{PH}^{obs}| - |(|F_P|.e^{i\alpha_P} + \underline{F}_H)|)^2$$

where α_P is the current estimation of protein phase. More recently, Dickerson, Weinzierl and Palmer (1968) have expanded this method to include the refinement of j derivatives alternating with the phase calculation for each of the h reflections. The refinement cycle treats each derivative independently. The function minimised is

$$E_j = \sum_h w_h.(|F_{(h,j)}| - |D_{(h,j(\alpha_P))}|)^2$$

where

$$D_{h,j(\alpha_P)} = K.(F_P^2 + F_{Hj}^2 + 2.\ |F_P|.|F_{Hj}|.\cos(\alpha_{Hj} - \alpha_P))^{\frac{1}{2}}$$

and w_h is the weight of the h th. reflection and is taken as

$w_h = 1/E_h$ where E_h is an error associated with the phase determination (see below).

Phase Determination.

The basis of calculating the phase is set down in a paper by Blow and Crick (1959), dealing with the errors involved in the multiple isomorphous replacement method. Fig. I. 4 must be looked on as free from all error. The Blow-Crick treatment assumes that errors arising from lack of isomorphism, incompletely refined heavy atom sites, and intensity measurements all can be considered as attached to the amplitude of $|F_{PH}|$ causing a lack of closure of the vector triangle, (Fig. I. 5). The error, $e(\alpha_P)$ for any reflection can be expressed, then, as

$$e(\alpha_P) = |F_{PH}| - |D(\alpha_P)|$$

where $D(\alpha_P)$ is the third side of the triangle, given by

$$D^2(\alpha_P) = F_P^2 + F_H^2 + 2.|F_P|.|F_H|.\cos(\alpha_H - \alpha_P)$$

For a derivative, the probability of the phase angle, α_P, being correct can be expressed as

$$p(\alpha_P) = N.\exp(-e(\alpha_P)^2/2.E^2$$

if it is assumed that the errors obey a Gaussian distribution and where N is a normalisation factor such that

$$\int_0^{2\pi} p(\alpha).d\alpha = 1$$

E is the root mean square error in $|F_{PH}|$ which may be found approximately from the centric data by the relation

$$\langle E^2 \rangle = \sum_{hkl} (|F_{PH}| - (|F_P + F_H|))^2/n.$$

For j derivatives, the probability is the product of the individual

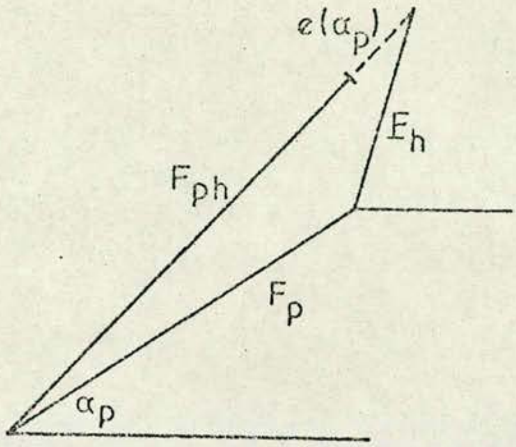


FIG. I.5 The lack of closure error shown in the $|F_{ph}|$ amplitude.

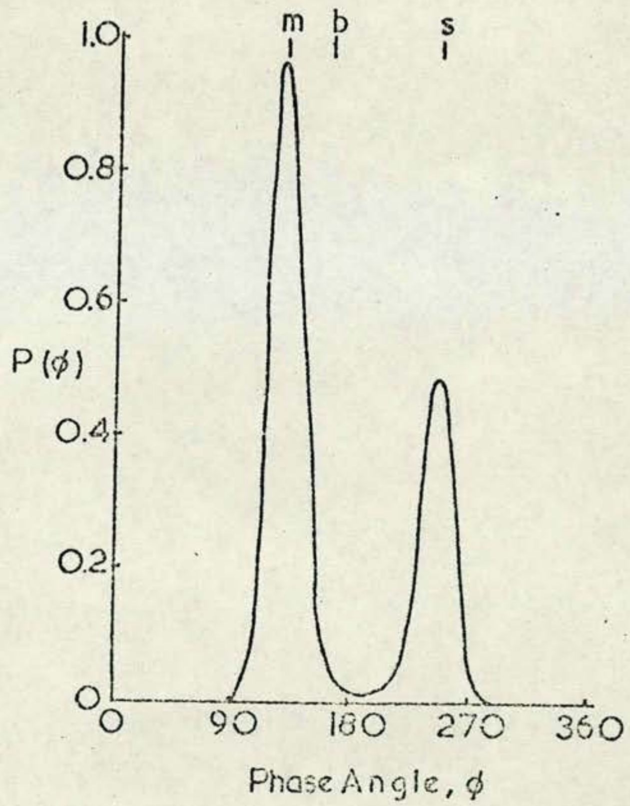


FIG. I.6 The probability $P(\phi)$ as a function of ϕ showing the most probable and best phases for a typical reflection.

probabilities. Thus,

$$p(\alpha_p) = N.\exp\left(-\sum_j (e_j(\alpha_p)^2/2.E_j^2)\right)$$

If this probability is plotted out against varying α_p, as in Fig. I. 6, for a typical reflection, it can be seen that the most probable value of α_p is not necessarily the best. Blow and Crick suggest that the centroid of this distribution gives the "best" Fourier since such a synthesis will give the minimum deviation between the true electron density and that calculated using the phases of the centroids of the probability distributions and appropriately weighted amplitudes.

The weighting scheme can be derived as follows. The mean square error in electron density contributed by one reflection and its conjugate can be written as

$$\langle\Delta\rho\rangle^2 = 2/V^2(\underline{F}_u - \underline{F}_t)^2$$

where $\underline{F}_u$ is the value used in the synthesis and $\underline{F}_t$ is its true value (Dickerson, Kendrew and Strandberg (1961)). Now, $\underline{F}_t$ can be supposed to have a magnitude $|F_p|$ and a probability $p(\alpha_p)$ of having a phase, α_p. If $\underline{r}(\alpha)$ is a unit vector in the direction α_p, then

$$\underline{F}_t = |F_p| .\underline{r}(\alpha)$$

and if this is substituted in the equation above

$$\langle\Delta\rho\rangle^2 = 2/V^2\int(\underline{F}_u - |F_p| .\underline{r}(\alpha))^2.p(\alpha_p).\ d\underline{r}(\alpha).$$

Minimising this with respect to $\underline{r}(\alpha)$ gives

$$\frac{d\langle\Delta\rho\rangle^2}{d\underline{r}(\alpha)} = \frac{-4\,|Fp|}{V2} \int (\underline{F}_u - |F_p|\ .\underline{r}(\alpha))p(\alpha_p).d\underline{r}(\alpha) = 0$$

Thus,

$$\underline{F}_{u(\text{best})} = \int |F_P| \,.\underline{r}(\alpha).p(\alpha_p).d\underline{r}(\alpha) \;/\int p(\alpha_p).d\underline{r}(\alpha)$$

$$= m.\,|F_P|\,.\exp i\alpha_{best}$$

where m is the mean value of the cosine of the error in the phase angle. In other words, m, the "figure of merit" is a measure of the reliability of the phase angle. In Fig. I. 7, this is shown diagramatically. The circle is of radius $1/|F_P|$, the vector $\underline{m}$ joins the origin to the centroid of the probability distribution, shown drawn round the circumference. $\underline{m}$ is represented by (m,α_{best}) and Dickerson et al. have shown that the computation of the best phase angle can be done by working out the probability at small intervals from 0^o to 360^o using the relationship on page 23 for the probability. That is

$$m.\cos\alpha_{best} = \sum_i P(\alpha_i).\cos(\alpha_i)/\sum_i P(\alpha_i)$$

and

$$m.\sin\alpha_{best} = \sum_i P(\alpha_i).\sin(\alpha_i)/\sum_i P(\alpha_i).$$

Finally, they have also shown that the error in the electron density map is given by

$$\langle\Delta\rho^2\rangle = \frac{2}{V^2}\sum_{\underline{h}} F_{\underline{h}}^2\,(1-m_{\underline{h}}^2)$$

In the centrosymmetric case, the phase is either 0 or π and the probability can be expressed more simply. The closure error, e , is written as

$$e = (\underline{F}_H \pm |F_{PH}| \mp |F_P|)$$

giving four possibilities, two that the sign of $|F_P|$ is positive,

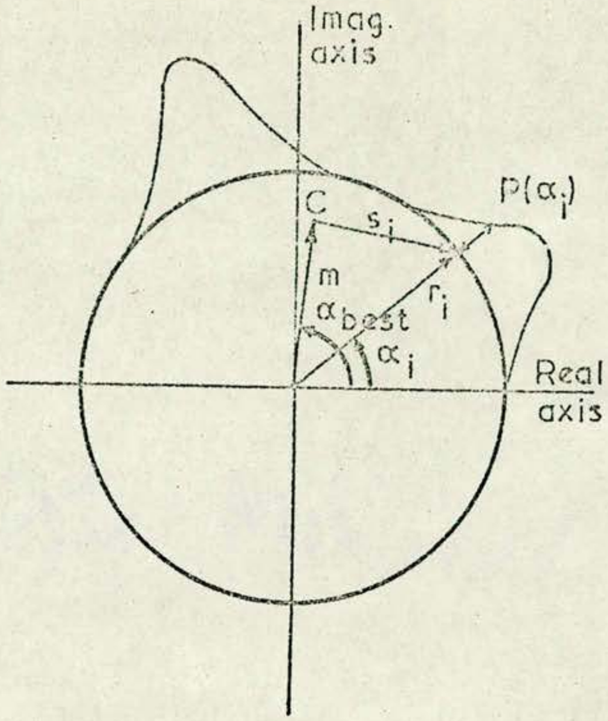


FIG. I.7 The figure of merit depicted as a vector to the centre of gravity of the probability distribution. (After Dickerson et al)

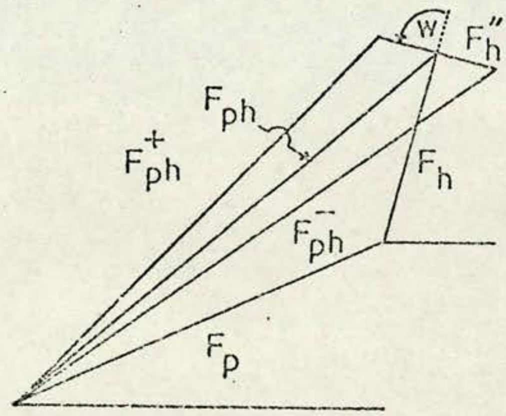


FIG. I.8 Diagram showing the anomalous scattering components of the hkl and $\bar{h}\bar{k}\bar{l}$ reflections.

two that it is negative. Thus,

$$P^+ = N.\left\{ \exp \frac{-(\underline{F}_H - |F_{PH}| + |F_P|)^2}{2.E^2} + \exp \frac{-(\underline{F}_H + |F_{PH}| + |F_P|)^2}{2E^2} \right\}$$

and

$$P^- = N.\left\{ \exp \frac{-(\underline{F}_H - |F_{PH}| - |F_P|)^2}{2.E^2} + \exp \frac{-(\underline{F}_H + |F_{PH}| - |F_P|)^2}{2E^2} \right\}$$

where N is the normalisation factor such that $P^+ + P^- = 1$, and P^+, P^- are the probabilities that the sign will be positive and negative respectively.

The Inclusion of Anomalous Scattering Effects.

Bijvoet (1954) has discussed the use of anomalous scattering to solve the phase ambiguity in the single isomorphous replacement method. In protein crystals, the only significant anomalous scattering with Cu radiation comes from the heavy atoms in metallo-proteins or from those included as in the isomorphous replacement method. The scattering factor for each atom, j, used in the calculation of the vector, $\underline{F}_H$, can be written as

$$f_j = f^o_j + \Delta f'_j + i.\Delta f''_j$$

$$= f'_j + i.f''_j \qquad \text{(James (1954))}$$

where f" is the imaginary part of the vector and is advanced relative to f'_j by $\pi/2$. A summation over all j anomalous scatterers whose imaginary parts can be different, gives the vector $\underline{F}_H$ made up of

$$\underline{F}'_H + \underline{F}''_H$$

with $\underline{F}''_H$ in advance of $\underline{F}'_H$. Thus the anomalous contribution can be added to the $\underline{F}_H$ vector with a phase advanced by w. This is shown in Fig. I. 8. North (1965) has treated the problem in the following manner. All of the anomalous scatterers are assumed to be the same (an assumption valid if, as is usually the case, only a single heavy atom complex is used for the derivative's preparation (Rossman (1961)) so that $w = \frac{\pi}{2}$ or $k = F'_H/F''_H$. If F^+_{PH}, F^-_{PH} refer to the values of $|F_{PH}|$ from the hkl and $\overline{hkl}$ reflections respectively, then from the diagram

$$F^{+\,2}_{PH} = F_{PH}^{\,2} + F_H^{''2} - 2.F''_H\,|F_{PH}|.\sin\Upsilon$$

and

$$F^{-\,2}_{PH} = F_{PH}^{\,2} + F_H^{''2} + 2.F''_H\,|F_{PH}|.\sin\Upsilon$$

so that

$$F^{+}_{PH} - F^{-}_{PH} = \Delta_{PH} = -2.F''_H.\sin\Upsilon$$

Thus, an indication of which phase of the two can be obtained from an examination of the sign of Δ_{PH}. An examination of the hkl and $\overline{hkl}$ reflections in a centrosymmetric zone from one crystal showed North that the root mean square errors associated with the breakdown of Friedel's Law were about a quarter of those for the heavy atom differences. He therefore proposed a similar treatment of the anomalous contribution to that of Blow and Crick so that the data could be included to increase the figures of merit. If $x'(\alpha)$ is the anomalous error then

$$x'(\alpha) = \Delta_{PH} + 2.F''_H.\sin\Upsilon$$

But $\sin\Upsilon = (|F_P|/|F_{PH}|).\sin(\alpha_H - \alpha_P)$ so that

$$x'(\alpha) = \Delta_{PH} + (2.|F_P|.F''_H/|F_{PH}|).\sin(\alpha_H - \alpha_P)$$

and $x'(\alpha)$ and E' can then be combined to give

$$P_{an}(\alpha) = N.\exp(-x'(\alpha)^2/2.E'^2)$$

The ambiguity in North's paper is which value of $|F_{PH}|$ should be used since it can either be found approximately from $\frac{1}{2}.(F^{+}_{PH}+F^{-}_{PH})$ or calculated from

$$F_c^2 = F_P^2 + F_H^2 + 2.\ |F_P|\,.\,|F_H|\,.\ \cos(\alpha_P - \alpha_H).$$

Matthews (1966a) has examined this ambiguity and finds that F_{PH}^{calc} is better. He also examined the inclusion of different anomalous scatterers arriving at the result

$$P_{an}(\alpha) = N.\exp(-1/2.E'^2).\ (-\Delta_{PH} + 2.\ F_H''.\cos(\Upsilon+w))^2).$$

F_H'' is found usually from the relationship

$$k.F_H'' = F_H'$$

where

$$k = 2.\ (\overline{|F_{PH} - F_P|})/(\overline{|F^{+}_{PH} - F^{-}_{PH}|})\ \text{Matthews (1966b).}$$

Difference Fouriers.

Having collected data for two derivatives which may not be isomorphous to resolutions much greater than 6Å, it is possible to produce phases which can be used to screen further derivatives more readily than the Patterson methods described above. Steinrauf (1963) has suggested using the following coefficients

$$(|F_{PH}| - |F_P|).\exp(i.\alpha_P)$$

This will show not only the heavy atom site but may also show up any minor ones so far undetected. It is weighted in favour of the terms where $|F_{PH}| - |F_P|$ is large so that the three vectors are nearly collinear.

Another possibility is to remove the major heavy atom site by calculating a "double-difference Fourier". The terms of this are

$$(|F_{PH}| - |\underline{F}_P + \underline{F}_H|).\exp(i.\alpha_{PH})$$

This could show up any undetected minor sites even more clearly than the difference Fourier.

CHAPTER 2

ALDOLASE: Introduction and previous work

The Reaction Catalysed.

Aldehydes, in the presence of a small quantity of alkali, can condense to form 'aldols'. This base catalysed, nucleophilic reaction requires an activated α-carbon on one of the aldehydes and can be summarised as shown in Fig. II. 1. If the two aldehydes (or ketone and aldehyde) are different, it can be seen that the reaction will give a mixture of all possible products. In living systems this type of reaction occurs a great deal and catalysts have evolved which can do it, and its reverse, with mixed starting materials and in the mild conditions prevailing in the living cell. These catalysts are known trivially as the aldolases or, more formally, as ketose phosphate aldehyde lyases. They appear in most living systems in one form or another but perhaps the most well known of these is that in the metabolic breakdown of glycogen to lactic acid taking place in muscle. The glycolytic pathway showing this aldol condensation in context, is shown in Fig. II. 2. Muscle fructose-1,6-diphosphate aldolase, E.C. 4.1.2.13, has been widely studied in recent years, most of the work being done on that isolated from rabbit.

Crystallisation

Rabbit muscle aldolase was first studied in any detail by Meyerhof and Lohmann (1934a & b) although crystals were not obtained for some five to ten years when Baranowski crystallised myogen A (1939) and then Warburg and Christian (1943) obtained crystals of aldolase from rat muscle. The rabbit enzyme was first crystallised by Taylor, Green and Cori (1948) and, since then, has been obtained by several methods

The 'Aldol' Reaction.

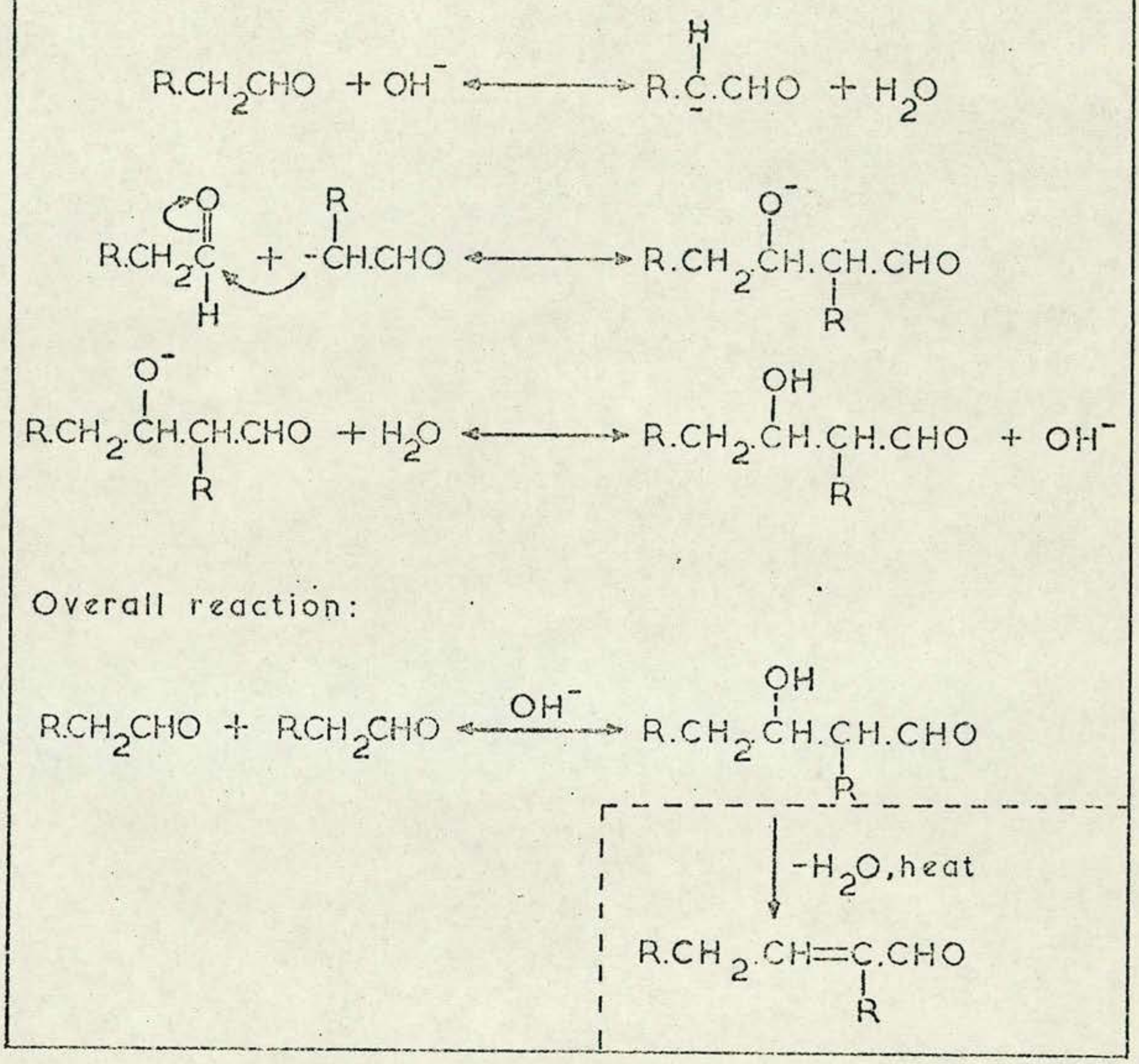
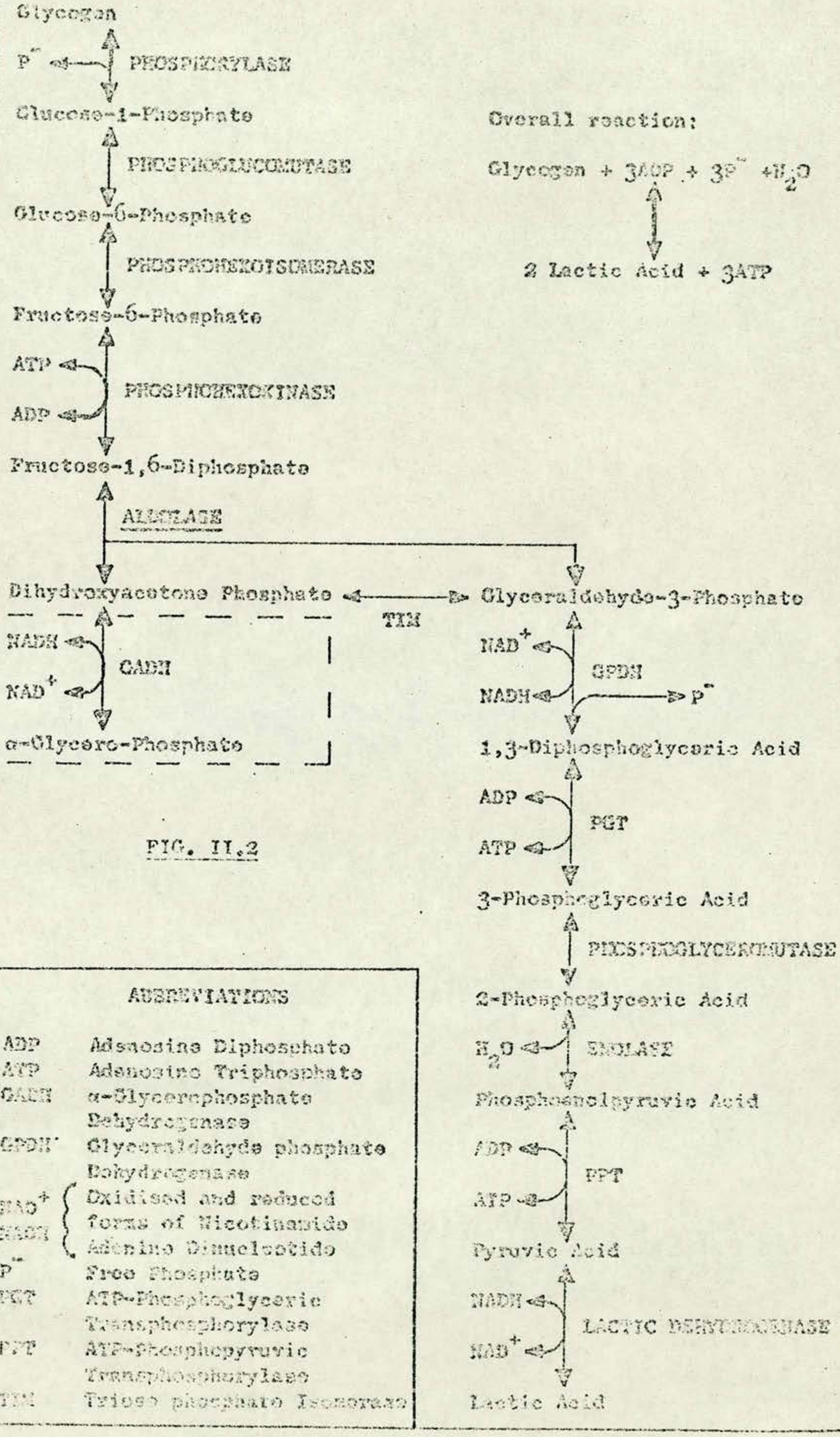


FIG. II.1

THE GLYCOLYTIC PATHWAY IN MUSCLE.



Overall reaction:

$$\text{Glycogen} + 3ADP + 3P^{-} + H_2O \rightleftharpoons 2\ \text{Lactic Acid} + 3ATP$$

FIG. II.2

ABBREVIATIONS

ADP	Adenosine Diphosphate
ATP	Adenosine Triphosphate
GADH	α-Glycerophosphate Dehydrogenase
GPDH	Glyceraldehyde phosphate Dehydrogenase
NAD^+, NADH	Oxidised and reduced forms of Nicotinamide Adenine Dinucleotide
P^-	Free Phosphate
PGT	ATP-Phosphoglyceric Transphosphorylase
PPT	ATP-Phosphopyruvic Transphosphorylase
TIM	Triose phosphate Isomerase

(Beisenherz et al. (1953), Wolf and Leuthardt (1957), Czok and Bücher (1960) and Eagles et al. (1969)). It crystallises in three forms, two of which seem capable of growing to appreciable size. The low pH form occurs at pH's less than 7 and is hexagonal bipyramidal (see the papers listed above). At pH values above 7.0, in ammonium sulphate solution, hexagonal plates are formed, these being first observed by Taylor et al. and being grown to a suitable size for X-ray work by McMurray and Gutfreund (Eagles et al.). Wolf and Leuthardt have observed that on crystallisation from A.S. the first crystals are invariably fine needles which will convert readily to hexagonal bipyramids at high protein concentration and low temperature. In sufficiently dilute solution, they claim, the needle form is the stable one, at least between pH 5.8 and pH 7.5 and at temperatures down to -10^{o}C.

Number of Sub-units

The subunit structure has been a controversial point until recently. The difficulty has been in deciding whether there are three or four subunits in the molecule, mainly because there has been a moderate amount of evidence to support both points of view. The dilemma has now been resolved in favour of four subunits but it is interesting to look at the evidence for there being three.

The evidence for this comes from several independent sources, namely end group analysis, the number of active sites, gel disc electrophoresis and ultracentrifugation.

Drechsler, Boyer and Kowalsky (1959) and Kowalsky and Boyer (1960) have examined the action of carboxypeptidase on aldolase. They found that three moles of tyrosine were released per mole of enzyme from the carboxy termini of the chains. Winstead and Wold (1964) confirm this result although all three papers admit that this value should be taken as a minimum.

The amino terminal analysis, as reported by Edelstein and Schachman (1966) and Sine and Hass (1967) gave three moles of proline per mole of aldolase. Two moles of proline have been obtained by Udenfreund and Velick (1951) using p-iodophenylsulphomyl ('pipsyl') chloride as the N-terminal "specific" reagent and, with 1,4-dinitro-2-fluoro benzene, Signor et al. (1969) found the same number.

The number of active sites has been reported variously as one, two and three. Westhead, Butler and Boyer (1963) found only one active site per molecule by means of equilibrium dialysis and ultracentrifugation with FDP and DAP. Most reports however, find three sites.

By using labelled DAP, Grazi et al. (1962) have been able to show that sodium borohydride reduction of the enzyme-substrate complex followed by complete acid hydrolysis yields β-glyceryllysine. Initially, Lai et al. (1965) found two active sites but Kobashi, Lai and Horecker (1966), on re-examination, have found that only three moles of radioactive β-glyceryllysine are produced per mole of enzyme. Binding studies of substrate analogues (Hartman and Barker (1965), Castellino and Barker (1966)) and of phosphate ions (Ginsburg (1966) and Ginsburg and Mehler (1966)) by competitive inhibition and equilibrium dialysis also show that there are three sites.

Further evidence came from Horecker (1966) and Chan, Morse and Horecker (1967). They found that dissociation of aldolase in 8M urea, pH 8.5 followed by gel disc electrophoresis gave two distinct bands, the densitometer trace of which showed a ratio of 2:1. This has also been observed by Edelstein and Schachmann (1966).

Finally, ultracentrifuge results of the dissociation of the enzyme in a variety of media seemed to point to three subunits. Deal, Rutter and van Holde (1963) studied the acid dissociation and found that it was completely reversible yielding at low pH, three highly unfolded chains.

Stellwagen and Schachman (1962) found three subunits by dissociation with HCl, acetic acid, urea and sodium dodecyl sulphate. These changes were reversible also. Schachman and Edelstein (1966) found three chains with guanidine hydrochloride and, with succinic anhydride at neutral pH, three subunits are again reported by Hass (1964). However, in alkaline media, six were found by Hass (1964) and Sine and Hass (1967) adding further confusion to already apparently conflicting reports.

It has now been established that, in fact, there are four subunits. This has been done in the following ways: ultracentrifugation, end-group analysis, specific cleavage, binding of substrate, hybridisation studies, electron microscopy and, most recently, X-ray analysis.

Kawahara and Tanford (1966) found there to be two pairs of subunits with molecular weights between 36,500 and 43,000 by ultracentrifugation in guanidine hydrochloride, which was also 0.1M in β-mercapto-ethanol, both at neutral and low pH's. Sia and Horecker (1968a & b) used maleic anhydride for the dissociation and have found four subunits of molecular weight about 44,300 whilst in acid solution they found four of molecular weight 40,000.

A re-examination of the end-group analysis has also found there to be four C -terminal tyrosine residues. Morse, Chan and Horecker (1967) digested S-carboxymethylated and partially dissociated (4M urea) aldolase with carboxypeptidase and found that four moles per mole of enzyme of tyrosine and alanine were released after 50 and 87 hours respectively. Sine and Hass, as reported in the paper by Morse et al., have achieved similar results with the native enzyme. More recently, partial sequencing of the C-terminal chains reveals two similar hexapeptides in equal quantities as shown below (Lai, Chen and Horecker (1970).

α : Ile-Ser-Asn-His-Ala-Tyr.

β : Ile-Ser-Asp-His-Ala-Tyr.

This, it is thought, may be the only difference, especially since Anderson and Perham (1970) find that there are three cysteines per subunit whose location and reactivity are identical throughout the four subunits. But isoelectrofocusing experiments show there to be an average structure of α_2-β_2 (Susor, Kochman & Rutter (1969)). Lai et al. also state that the Asp-His linkage has been found to be very stable to carboxypeptidase digestion and that this could be the explanation for the earlier attempts only giving three tyrosines per mole. The N-terminal analysis carried out by Hass (1964) states that there are at least four prolines per mole.

Cleavage of the molecule with CNBr, a reagent specific for the peptide bond next to a methionine residue, has been shown to give four identical peptides of 41 residues each, around the active site (Lai and Chen (1968)).

Lai et al. (1968) have shown that GAP and erythrose-4-phosphate cause highly specific, irreversible inactivation of aldolase on binding four moles to the enzyme.

Some very compelling evidence derives from the elegant hybridisation studies of Rutter and his co-workers (Penhoet et al. (1966) and (1967)). Various parts of rabbit (and,indeed, of all animals) contain different forms of aldolase distinguished by differing FDP activities, electrophoretic mobilities, and sequences (see, for example, Lebherz and Rutter (1969) and Penhoet et al. (1969)). These forms have been labelled A,B and C for the iso-enzymes isolated from muscle, liver and brain respectively, although the brain usually appears to have a mixture of A and C. Now brain aldolase on gel electrophoresis at pH 8.6 yields five bands whilst muscle aldolase under the same conditions gives but one. Rutter has isolated the pure form from brain (C) and found that it, too, is electrophoretically homogeneous. Thus, it appeared that the five bands must be formed from a hybrid set with pure A and pure C

Author	Method	Molecular Weight	No. of Subunits	Subunit M.W.	Dissociating Medium
Taylor & Lowry (1956)	Ultracentrifugation	140,000	-	-	-
Kawahara & Tanford (1966)	"	158,000-161,000	4	36500-43000	6M Gu-HCl, 0.1M βME
			2	80000	Dilute solutions
Castellino & Barker (1968)	Membrane osmometry	156,000 ± 1000	3.7 - 3.95	42400 ±300	6M Gu-HCl, 0.1M βME
Sia & Horecker (1968a)	Ultracentrifugation	159,000 ± 2000	4	44300 ±1500	Maleic anhydride
Sia & Horecker (1968b)	"	162,500	4	40000 ±2000	Acid, βME
Sine & Hass (1969)	"	156,000-160,000	4	40200-41400	Borate, pH 12.5
Gracy et al. * (1969)	Gel filtration & u/cent	158,000	4	37900-39800	6M Gu-HCl, Na dodecyl sulphate

* This data is for the isozyme from liver.

TABLE II.1 Some recent molecular weight determinations on rabbit muscle aldolase.

at the extremes. In order to check this, the centre one of the set was dissociated at pH 2.0, reassociated and then treated as before whereupon a five membered set was obtained. Repeating the above experiments with an 'artificial' brain extract made from a radioactive sample of rat muscle enzyme (A) and unlabelled rabbit brain enzyme gave identical results and the radioactivities of the five bands were in the ratios 1:0.75:0.5:0.25:0. These data are consistent with a four subunit molecule for aldolase, hybridisation of which would give : A_4,A_3C,A_2C_2, AC_3,C_4. In fact, it is hard to see how a three subunit molecule could fulfill these requirements even if one 'pure' molecule is of the form $\alpha_2\beta$. Lai (1968) has found that the α- and β-subunits are almost identical from the tryptic fingerprints and Penhoet et al. (1969) have observed far more striking differences between forms A,B and C. Besides Rutter (Penhoet et al. (1966), (1967)), this phenomenon has also been observed in chicken (Herskovits et al. (1967)), rat (Christen et al. (1966)), guinea-pig (Nicholas and Bachelard (1969)) and again in rabbit (Foxwell, Cran and Baron (1966), and Guertler and Leuthardt (1970)).

Finally, electron microscopy has shown that the molecule is a tetramer with a subunit diameter of about 40Å (Penhoet et al. (1967)). Also, X-ray work by Johnson and her colleagues (Eagles et al. (1969)) and this work, shows there to be four subunits per molecule.

The Molecular Weight

The molecular weight has been measured by various techniques and has been quoted as having values between 142,000 (Stellwagen and Schachman (1962)) and 160,000 (Sia and Horecker (1968)). However, most measurements made recently seem to agree with that of Kawahara and Tanford (1966) of 158,000. These values and the methods used to obtain them are shown in Table II. 1.

Amino Acid	Velick & Ronzoni (1948).	Shimizu & Ozawa (1967).	Penhoet, Kochman & Rutter (1969).	Lai (1968).	Anderson, Gibbons & Perham (1969).
Lysine	103	100	121	104	108
Histidine	43	44	45	44	40
Arginine	58	59	67	56	60
Asparagine	115	119	112	116	116
Threonine	86	85	80	88	84
Serine	99	82	75	84	80
Glutamine	123	160	160	164	164
Proline	78	75	78	80	76
Glycine	118	123	116	124	120
Alanine	152	175	154	172	164
Cysteine	29	32	28	32	28
Valine	100	90	90	80	84
Methionine	12	13	12	12	12
Isoleucine	95	83	83	76	80
Leucine	139	142	140	140	132
Tyrosine	46	43	44	48	44
Phenylalanine	20	28	29	28	28
Tryptophan	18	13	12	12	16
	1443	1466	1446	1460	1436

TABLE II.2 Amino acid analyses for rabbit muscle aldolase. The data of Velick and Ronzoni have been corrected for a molecular weight of 158,000 and those of Anderson et al. have been rounded to the nearest whole number and then multiplied by four.

```
H2N-Pro-His------Cys-Ala-Gln-Tyr-Lys-Asp-Gly-Ala-Asp-Pho-Ala-Lys-Arg-Met-Glu-Asn-Ala-Asn-Val-Leu-
                                        150                155                 160

Ala-Arg-Tyr-Ala-Ser-Ile-Gly-Val-(Cys,Asx,3Glx,Pro,Ile,Leu)-Pro-Glx-Ile-Lou-Pro-Gly-(His,3Asx,Val)-
165               170                                      181               185

Lys-Arg-Cys-Gln-Tyr-Val-Thr-Glu-Lys-Val-Leu-Ala-Ala-Val-Lys-Ala-Leu-Sor-Asn-His-His-Ile-Tyr-Leu-
192         195                 200                 205                 210                 215

Gln-Gly-Thr-Leu-Leu-LYS-Pro-Asn-Mot-Val-Thr-Pro-Gly-His-Ala-Cys-Thr-Gly-Lys-Tyr-Ser-His-Glu-Gln-
                220                 225                 230                 235

Ile-Ala-Mot-Ala-Thr-Val-Pro-Ala------Ala-Leu-Ala-Cys-Ser-(Asn,Gln,Gly,Ala,Leu)-Lys------Ile-Ser-
240                  245

Asn-His-Ala-Tyr-COOH
```

FIG. II.3 Partial sequence of rabbit muscle aldolase from Lai and Chen (1968) with corrections by Gibbons, Anderson and Perham (1970) and Szajani et al. (1970) and the sequence from residue 144 to residue 206 that of Sajgo (1971), F.E.B.S. Letters, 12 340. The active-site lysine is in capitals and the asparagine in the antepenultimate position is that which changes to aspartic acid in the β form.

Amino-acid Analysis and Partial Sequence

The amino acid analysis, like the other properties, has also been reported by several authors. The first and the four most recent are shown in Table II. 2. Benesch, Lardy and Benesch (1955) have found that about 29 moles of silver bind to aldolase in the presence of urea. Swenson and Boyer (1957) have found that 28 moles of pCMBS bind to the enzyme and Eagles et al. (1969) have found that 28 moles DTNB bind, providing that some denaturant such as 4M urea is present. These values are consistent within themselves but are in slight disagreement with two of the analyses. Could one thiol per subunit be escaping detection by the titrimetric methods employed? The final answer must wait on the complete sequence or the three-dimensional structure, whichever is the first. It is also known from the CNBr cleavage studies of Lai (1968) and Lai and Chen (1968) that there are 12 methionines. These workers have also done some sequencing studies on the peptides produced by such cleavage. Their data are shown for both α and β chains in Fig. II. 3. Anderson, Gibbons and Perham (1969), however, whilst agreeing with the N-termini found by CNBr cleavage, found certain inconsistencies in the sequence as reported by Lai. Gibbons, Anderson and Perham (1970) have checked the sequence and their corrections are shown in Fig. II. 3 as well. Szajani et al. (1970) have partially sequenced the same region and agree with the work of Perham. Also, they have found that the active lysine is about the 230th residue in the chain, and that there are two small sections of peptide chain round the other two cysteines with sequences as shown at the foot of the Fig. II.3, the upper being found in the last 120 residues of the chain. Thus, the number of amino acid residues would appear to be about 1450, or about 360 per chain. In fact, Lai (1968) and Lai, Chen and Horecker (1970) wonder if the differences between the

α and β chains could be only in a few residues at the C- and N-termini, since work by Koida and Lai (1969) and Koida et al. (1969) shows that the amount of β-chain is age dependent in rabbits and that the two terminal polypeptides have the only differences.

Chemical Results Concerning the Active Site.

A picture can be built up of the active site of the enzyme from chemical evidence. To begin with, there are four subunits as has been seen and there are four active sites per molecule (Lai et al.(1968)), probably one per subunit. Of these, it is thought by Kobashi et al. (1966), one appears to be blocked by phosphate since one mole is found per mole of enzyme on analysis of the crystalline protein. This is the most probable explanation of the various reports of the number of active sites being three. Ginsburg (1966) and Ginsburg and Mehler (1966) have shown that there are two types of binding sites for phosphate, one strong one for the 1-phosphate of FDP, the other weaker for the 6-phosphate (Morse and Horecker (1968)) and also that no binding occurs in 4M urea indicating that an intact tertiary structure is necessary for binding and hence for full activity. Hartman and Barker (1965) have shown that binding of diphosphate substrate analogues is more efficient than that of monophosphates and also that, of the diphosphate substrates, those with the phosphate groups some 10-12Å apart bind most strongly.

As has been already mentioned, a lysine residue is involved with the DHAP moiety since reduction with borohydride gives a β-glyceryllysine residue on hydrolysis (Grazi et al. (1962 a & b), Speck et al. (1963)). However, it is still impossible to say which of the other residues on either side of the lysine (Fig. II. 3), if any, are involved in the mechanism except that there are several hydrophobic and aromatic residues in close proximity which are perhaps the "lining" of the active cleft so that the charges of the various side chains involved in the mechanism are not "diluted" or smeared out by polar interaction.

Drechsler, Kowalsky and Boyer (1959) have observed that removal of the C-terminal tyrosines causes a drop in the activity with respect to FDP whilst having little or no effect on the rate of cleavage of F-1-P. Richards and Rutter (1961), Rutter, Richards and Woodfin (1961) and Rose, O'Connell and Mehler (1965) have confirmed this but the suggestion that the terminal tyrosine is involved only in the binding of the 6-phosphate in FDP has been challenged by Spolter, Adelman and Weinhouse (1965) and Rose et al. (1965) who have shown that aldehydes can partially restore the activity of carboxypeptidase-treated aldolase whilst having no effect on the native. This indicates that the tyrosine might be responsible for the removal of a proton from the Schiff-base intermediate (see below) to form a carbanion for the aldolisation step. In fact, Mehler and Cusic (1967) have shown that the binding of various substrates, which is controlled by the binding of the phsophate groups in the first instance, is independent both of the chain length between phosphates for $C_5 \rightarrow C_8$ and of carboxypeptidase treatment, implying a certain amount of flexibility in the active site. Hoffee et al. (1967) have shown similar effects to the removal of tyrosine by photo-oxidation of about half of the histidine residues in the enzyme with rose bengal. This reduces the activity to about 15% of its original value but aldehydes, in particular erythrose-4-phosphate, can restore the activity. Also the C-terminal tyrosine is made more susceptible to iodination and acetylation which could mean that the terminal residue and a histidine are both involved in proton transfer.

Titration of the enzyme with DTNB shows that four cysteines react quickly, eight more slowly after which the activity is destroyed and a further 16 if unfolding is first carried out with urea or guanidine hydrochloride (Eagles et al. (1969)). Further, FDP protects the enzyme against attack with DTNB on four of the eight slow-reacting thiols.

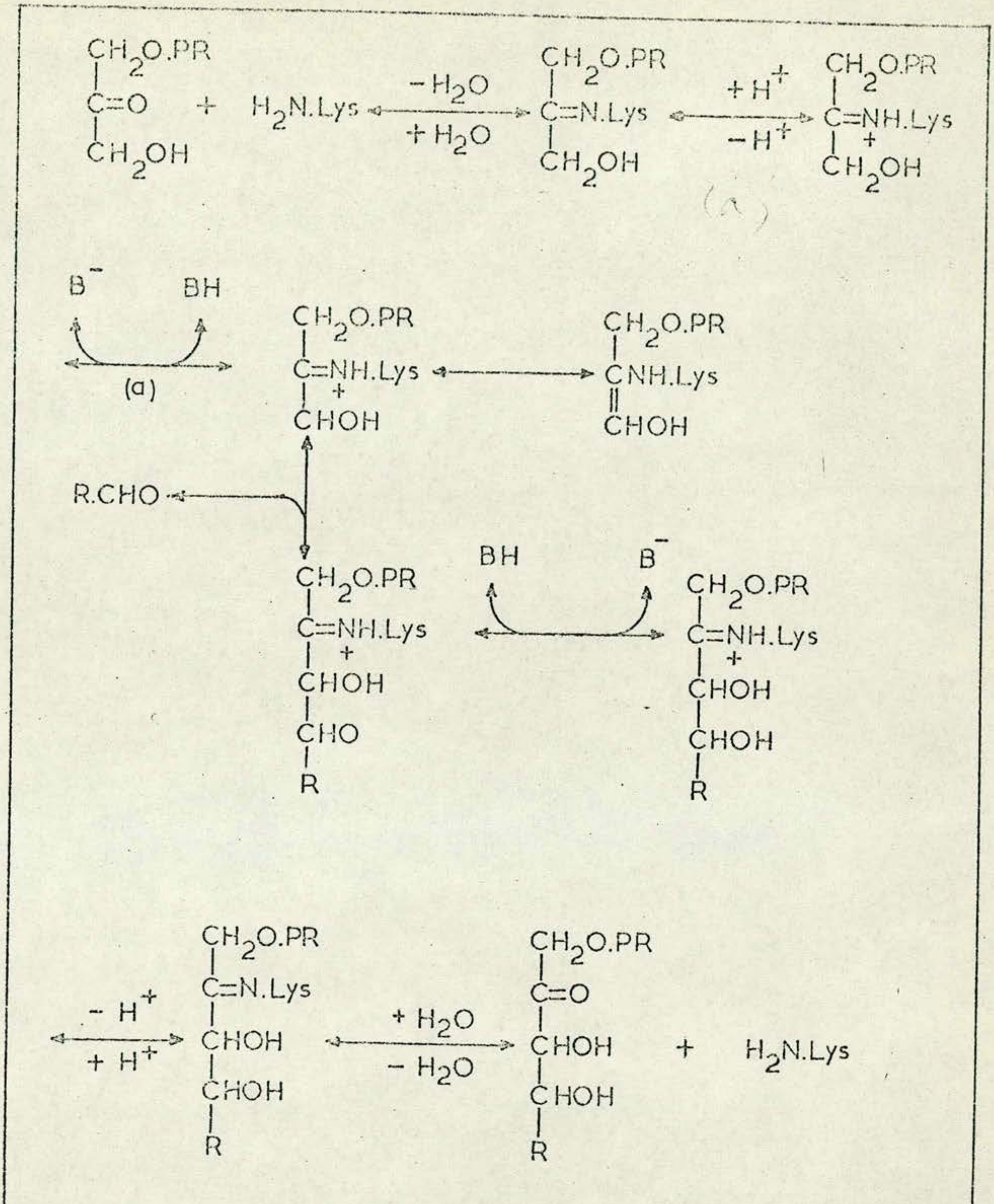


FIG. II.4 Mechanism of rabbit muscle aldolase proposed by Morse and Horecker (1968).

Swenson and Boyer (1957) and Szabolsci and Biszku (1961) have found similar results with pCMBS and so have Kowal, Cremona and Horecker (1965) with 2-chloro-1,4-dinitrobenzene. Lai and Hoffee (1966) have found that bromopyruvate, a similar compound to DHAP, and photo-oxidation seems to S-alkylate about six thiols causing complete inactivation. FDP, as above, will prevent alkylation of four of these. Thus, it would appear that there is a cysteine residue associated with the active site but not necessarily with the mechanism, and recently, Szajani et al. (1970) have confirmed this, finding that the active thiol group is in the last 120 residues of the chain.

In all of the experiments mentioned in the last section, the presence of the substrate has caused some measure of protection to the enzyme by slowing down or stopping the rate of activity loss. This is an indication of the firmness with which the substrate is held by the molecule and agrees with the theory of 'substrate-induced fit' put forward by Koshland (1963). In this, the substrate induces the correct spacing of the phosphate-binding sites by its binding into the active site which gives the enzyme the optimum conformation for the cleavage reaction and hence enhances the rate for this particular reaction.

Mechanism of Action

This last section has given a summary of the findings by chemical probing of the active site. A fuller account is contained in the review by Morse and Horecker (1968), who also give a proposed mechanism based on all the findings of the various techniques. This is shown in Fig. II. 4 where the importance of the Schiff-base forming lysine residue as a stabilising influence to the carbanion, can be seen. As mentioned above, it is the protonation/deprotonation, (a), step which is affected by the treatment with carboxypeptidase or photo-oxidation.

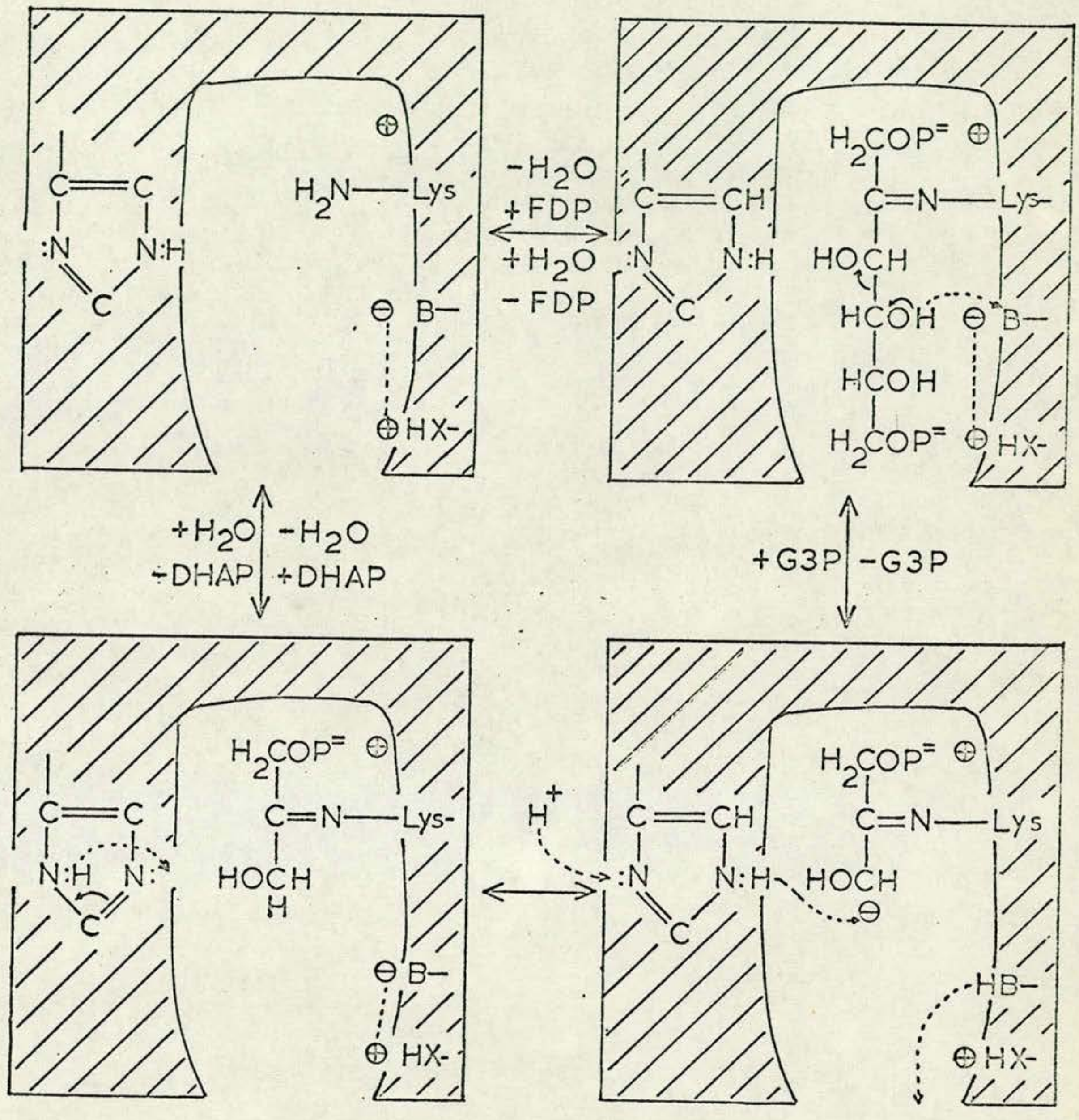


FIG. II.5 Proposed 'active site' region of rabbit muscle aldolase. Morse and Horecker (1968).

The basic group or groups (B) are not known as yet. Further, the mechanism has been 'fitted' by Morse and Horecker to a cleft of the type expected in enzymes since the work of Phillips and his colleagues on lysozyme (Blake et al. (1965)). This is shown in Fig. II. 5. The cleft is in an apolar region of the protein formed, not necessarily by adjacent residues, from the three-dimensional folding of the polypeptide chain and the extra stability gained by tetramerisation. The substrate is drawn into the cleft by means of the strong 1-phosphate binding site interacting with the phosphate group. The distribution of charge within the cleft is then such that the lysine residue can form a covalent bond with the DHAP/FDP. The second phosphate can then react with a second positive binding site which, it is suggested, might only become available on binding of the substrate at the other position. This latter suggestion, however, appears to be in conflict with the results obtained by Ginsburg and her coworkers on the binding of phosphate and sulphate ions, unless only a phosphate group is required to cause this change from neutral to positive. The basic group accepts a proton thereby initiating the dealdolisation step. The rest of the reaction then proceeds as shown giving eventually the starting structure again.

The detailed mechanism, the actual structure of the active site, the groups involved in binding and holding the substrate in the required conformation for cleavage or condensation are questions which a detailed molecular structure of the protein would go a long way towards answering. It was to this end that the following study was undertaken and, whilst it was realised that the complete structure would take a number of years, it was hoped to be able to find a form on which X-ray techniques would eventually lead to a solution. Some information about the gross molecular geometry and the number of subunits, which was still in doubt when the work was started, was also hoped for.

CHAPTER 3

ALDOLASE: Crystallisation

Before any study of a molecule as complex and as fragile as a protein is begun, it is necessary to have some means of estimating its quantity and quality. With enzymes, the obvious method is to follow the reaction catalysed and to define the activity as the number of micromoles of substrate transformed per minute (Enzyme Nomenclature (1965)). The specific activity is then the number of units per milligram of protein. By purifying the preparation until the specific activity increases no further, by which time it should be capable of crystallisation, it is possible to achieve a "pure" sample. The method of following the enzyme-substrate reaction varies from one protein to the next and there may even be a choice. Aldolase is a case where there are four possible techniques.

Choice of Assay.

The reaction catalysed by aldolase, as has already been seen, is

$$\text{FDP} \rightleftharpoons \text{GAP} + \text{DAP}$$

The equilibrium constant has been given by Meyerhof (1951) as 6.8×10^{-5} moles/l at pH 7.0 and 20°C so that some means of removing the products formed must be used. This then allows the products to be measured chemically by determining the alkali labile phosphate released in a given time (Herbert et al. (1940)) or by forming the 2,4-dinitrophenyl-hydrazones (2,4-DNPH) (Sibley and Lehninger (1949)). Alternatively, by making use of the metabolic pathway involving aldolase, the reaction can be coupled with other enzymes so that the well-known optical density change at 340nm. of the oxidation of NADH can be utilised. As large concentrations of phosphate were envisaged as a possible means of salting out the protein, the first method mentioned was discarded.

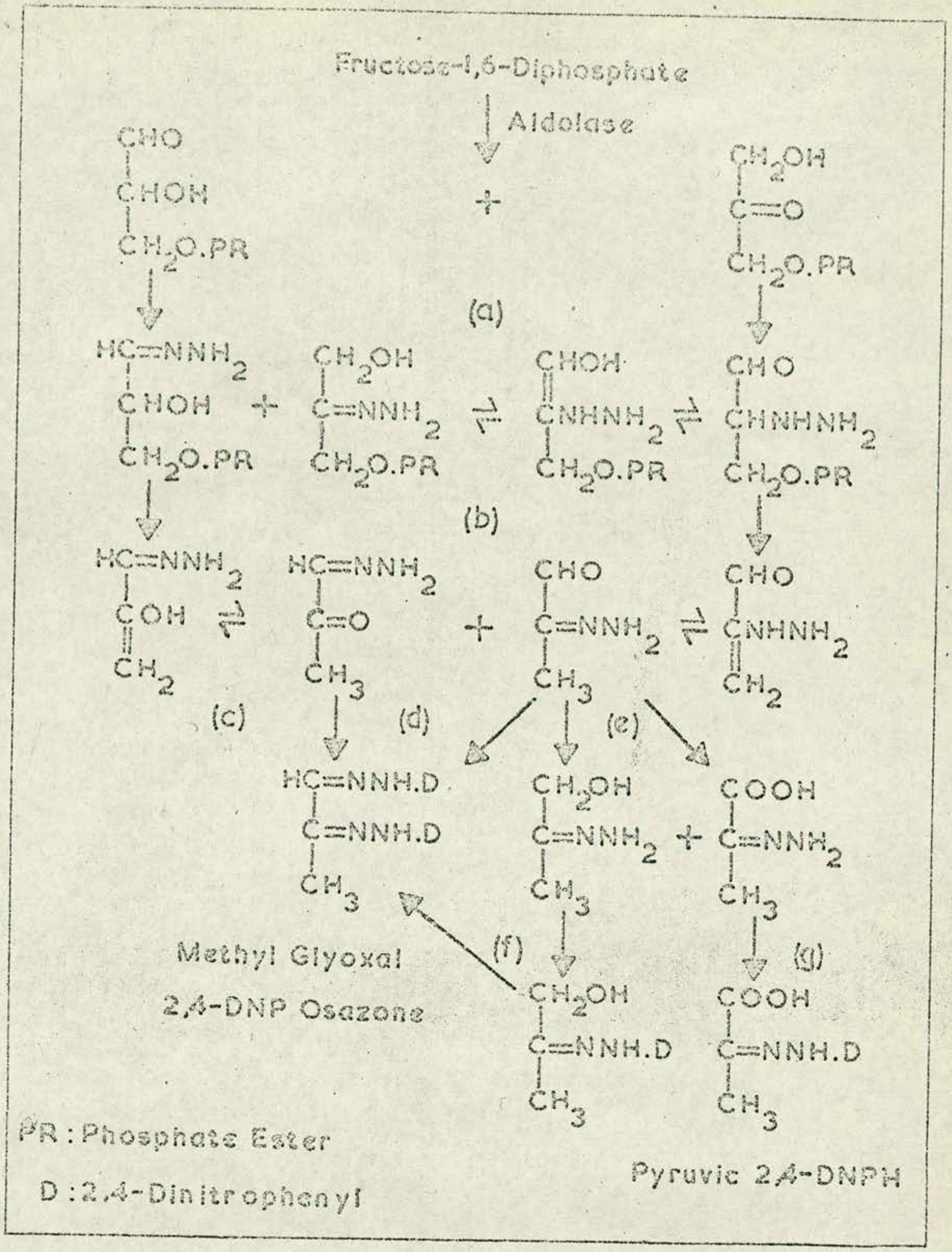


FIG. III.1 Proposed reactions involved in the 2,4-Dinitrophenyl hydrazine assay procedure.

Of the other two, the 2,4-DNPH one is independent of salt and other enzymes although it does take rather a long time. The coupling method, though unaffected by salt, is liable to interference from other systems present in a crude extract. It also requires several rather unstable solutions for its operation, but it is fast and accurate. Thus, during the early stages of a preparation the 2,4-DNPH method was used and when the purification had been nearly completed, during recrystallisations and when a series of assays was anticipated, the NADH method was used.

The Colorimetric Method.

This is the method of Sibley and Lehninger as modified by Bruns (1954). The triose phosphates produced by aldolase from FDP are trapped with hydrazine (Fig. III. 1, step (a)). The reaction is then stopped with trichloroacetic acid, the solution filtered to remove the deproteinised enzyme and then treated with excess NaOH to liberate the trioses, step (b). Treatment with 2,4-DNPH then produces a mixture of methylglyoxal-2,4-dinitrophenylosazone (MGDNP) and pyruvic acid-2,4-dinitrophenylhydrazone (PADNP) as shown. These are then dissolved in excess alkali producing dyes with absorption maxima at 540nm. Beck (1954) has observed that the ratio of extinction coefficients E_{dhap}/E_{gap} drops from 1.8 after 10 mins. to 1.0 after 60 mins. and ascribes this to the differing rates of formation of the osazone, present finally as 94% of the product, from DHAP and GAP. In the tautomerism shown in step (c) the hydrazine is not such a powerful 'driving force' as the carbonyl in step (d), which could explain Beck's result. Step (e) is the base-catalysed Cannizzaro reaction which produces pyruvic acid hydrazone, steps (d) and (g) are the exchange of hydrazine by 2,4-DNPH and step (f) is the

Procedure for the Colorimetric Assay

Pipette into clean dry test-tubes:

	Sample	Blank
Solution I	2.00mls.	2.00mls.
Enzyme solution	x mls. where x	1.00mls.
Distilled water	(1-x)mls.	-

Mix and incubate at 37°C for exactly 1hr. then add:

Trichloroacetic acid	3.00mls.	3.00mls.
Enzyme solution	-	x mls.
Distilled water	-	(1-x) mls.

Mix and filter (or centrifuge), then:

Filtrate	1.00mls.	1.00mls.
NaOH	1.00mls.	1.00mls.

Mix and allow hydrolysis to proceed for exactly 10 mins. at room temperature, then:

2,4-D NPH	1.00mls.	1.00mls.

Mix and incubate for exactly 10 mins. at 37°C then add:

NaOH	7.00mls.	7.00mls.

Pour these solutions, after mixing, into cuvettes and measure the optical density of sample versus blank at 540 nm.

oxidation of the alcohol to aldehyde followed by its conversion to the osazone. Beck's recommendation, therefore, is that timing of the chromophore development should be increased to 60 mins. However, providing that the timing is strictly observed reproducible results can be obtained, Bruns (1954). The solutions required are the following:

0.0048M collidine buffer pH 7.4

0.56M hydrazine sulphate

0.1M fructose-1,6-diphosphate (Na salt)

0.0016M sodium iodoacetate (to inhibit G-3-PDH)

0.1% 2,4-DNPH

plus 2M HCl, 0.75M NaOH and 10% w/v trichloroacetic acid. The first four of these solutions are mixed together in the ratios 100:25:25:25 plus 25 parts water, giving Solution I. The procedure for the assay is shown opposite. As can be seen, it is rather tedious.

<u>The Dynamic Assay.</u>

There are, in fact, two possible coupling reactions involving aldolase. The first, used by Warburg and Christian (1943) and Baranowski and Niederland (1949), uses the enzyme glyceraldehyde-3-phosphate dehydrogenase to produce 1,3-diphosphoglycerate. This utilises one mole of NADH per mole of FDP and solutions must contain <u>no</u> GDH/TIM. The other method is that of Racker (1947) which couples two enzymes to the reduction to α-glycerophosphate of DHAP. The reactions are shown below, and in the dashed box in Fig. II. 2.

$$\text{DAP} + \text{NADH} \underset{\text{GDH}}{\rightleftharpoons} \text{GLP} + \text{NAD}^+ + \text{H}^+ \quad K=1.78\text{x}10^4 \text{ moles/l. pH7.0, } 25^\circ\text{C}$$

and

$$\text{GAP} \underset{94\%}{\overset{\text{TIM}}{\rightleftharpoons}} \text{DAP} \qquad K=22 \text{ at pH7.0,} 25^\circ\text{C}$$

This gives an overall reaction

Procedure for the Dynamic Assay

Pipette into clean, dry test-tubes:

	Sample	Blank
Solution I	2.50mls.	–
NADH solution	0.05mls.	–
GDH/TIM suspension	0.01mls.	–
Phosphate buffer	(0.2 -x)mls.	(2.76 -x)mls.

Incubate at 37^{o}C for a minute to remove any DAP or GAP impurity and transfer to cuvettes in a thermostatted spectrophotometer.

Enzyme solution	x mls.	x mls.

In practice, it was found that with semi-pure and pure samples of the enzyme, x was so small that the reference was usually phosphate buffer, since the difference was being measured as a function of time. The enzyme is added to the blank first and then to the sample cuvette by means of a stirrer, a stop clock being started at the same time. The optical density is noted down every 30 secs. until a linear trend is apparent, usually 10 mins.

$$FDP \rightleftharpoons 2DHAP + 2NADH \rightleftharpoons 2GLP + 2NAD^{+} \quad K=4.75\times10^{5}(moles/l)^{-1}$$
$$pH\ 7.0, 25^{o}C.$$

and, if the rate determining step is made the breakdown of FDP, that is, on the amount of aldolase present, this provides a better method of assay than the first because it utilises two moles of NADH per mole of substrate, the rate of decrease of the NADH peak at 340nm. being a magnified measure of the enzyme's activity. Iodoacetate is added to inhibit the action of G-3-PDH (Baranowski and Niederland (1949)). The solutions required for the Bruns and Bergmeyer (1965) modification of the above assay are:

0.112M collidine buffer, pH 7.4

0.0012M iodoacetate (Na salt)

0.012M FDP (Na salt)

0.02M NADH (25mg. in 2mls. 1% $NaHCO_3$)

A suspension of 2 mg./ml. of GDH/TIM containing 1.8mg GDH and 0.2mg TIM in 2.4M A.S.

0.9% NaCl as reference

The first three of these solutions are mixed together in ratios 50: 25:25 plus 25 parts water giving Solution I. The procedure is shown opposite.

The solutions for both the assays described can be obtained in kit from C.F. Boehringer & Soehne Gmbh., Mannheim and in fact it was these kit procedures which were used.

The spectrophotometer available for this work was a manual recording Unicam SP500 Series I with a constant temperature cell housing thermostatted to the incubation temperature, 37°C.

Calibration of the Colorimetric Assay.

The method of calibrating this assay requires accurate analysis of minute quantities of alkali-labile phosphate produced on their splitting off from the triose molecules. As this method was considered

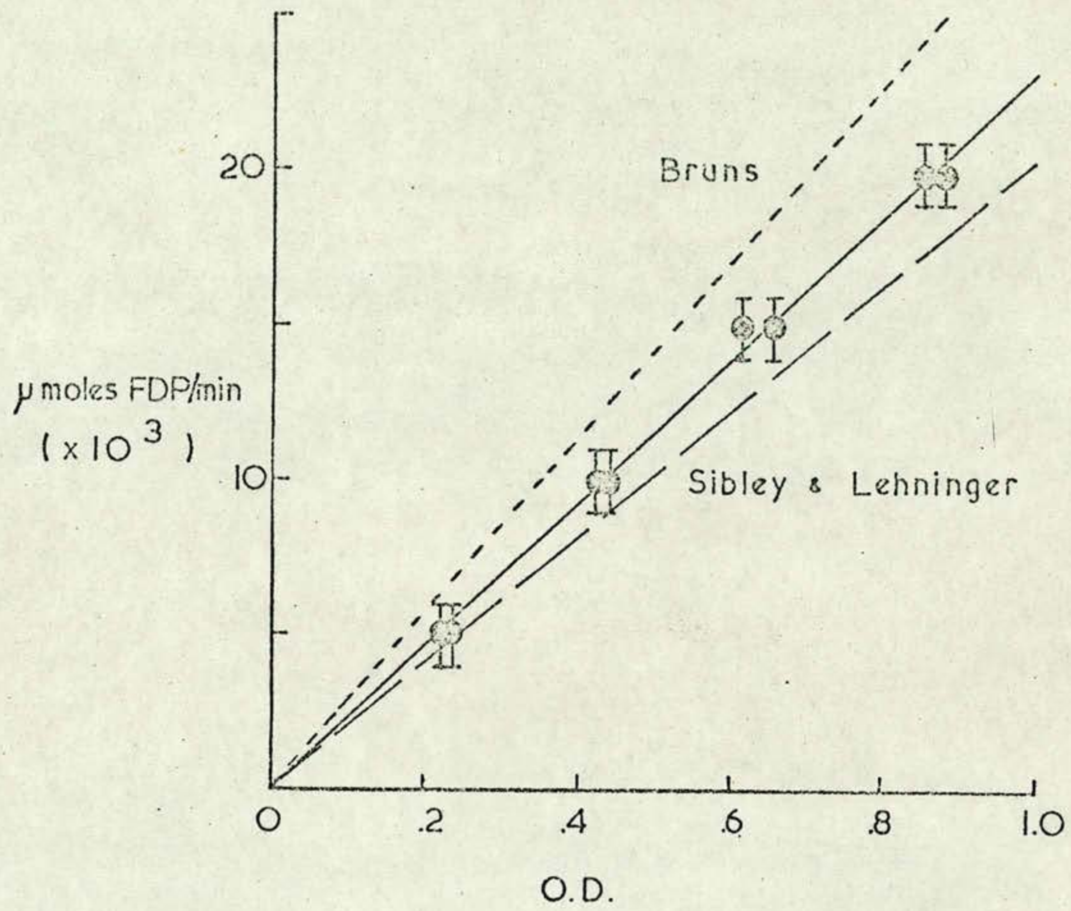


FIG. III,2 Calibration curve for the colorimetric assay.

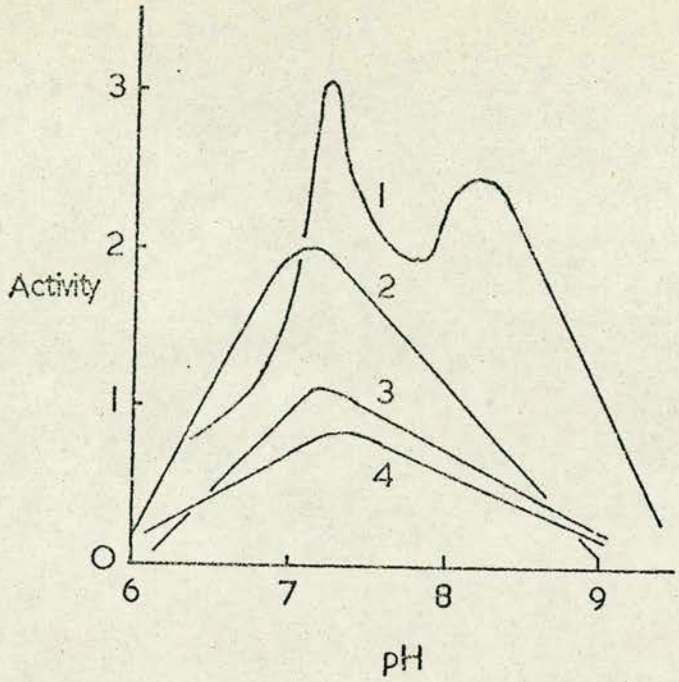


FIG. III.3 pH profiles of specific activity in various buffers after Dounce, Barnett and Boyer (1950). The activity is in arbitrary units. 1: Tris-HCl, 2: collidine, 3: nothing, 4: phosphate.

Buffer	pH Maximum	Activity relative to that in PO_4^{3-} at pH 7.4	
None*	7.2	1.48	
Phosphate*	7.4	1.0	
Collidine*	7.2	2.2*	1.7**
Tris*	7.2 & 8.2	3.0	
Veronal**	8.5	1.2	
Borate**	6.5	0.6	

* Data from Dounce et al. (1950) ; see also Fig. III.3.
** Data from Bruns and Bergmeyer (1965).

TABLE III.1 Data showing the effect of buffer on the activity of rabbit muscle aldolase.

to be rather less accurate than the dynamic assay on account of its many stages and, since it was only to be used as guide in the initial separations, it was thought adequate to accept the value given by Boehringer for their commercial preparation and use this for the calibration. Both Sibley and Lehninger and Bruns (1956) show that the reaction as performed is linear to optical densities above 1.0. Fig. III. 2 shows the published calibrations along with that of this work. The two published methods differ only in the buffer used and the pH, which has been chosen at the maximum rate of cleavage of the FDP for that buffer. It can be seen that the values differ quite markedly and this it has been observed is a function of the enzyme's activity changing with its solvation (Taylor 1957)). The data shown in Table III. 1 and Fig. III. 3 serve to illustrate this point. Taylor, Richards & Rutter (1961), and Bernfield, Berkeley & Bieber (1965) also note that full activity can only be preserved on dilution if some protective protein is added, for example, serum albumin.

The activity is then the dilution times the graph reading.

<u>Calibration of the Dynamic Assay.</u>

The rate of oxidation of NADH is something which can be measured quantitatively by using its molar extinction, $E_o = 6.22 \times 10^3$ and following the rate of disappearance of the peak at 340nm. This can then be related directly to the rate of FDP cleavage and hence to the activity of the enzyme.

Activity = $\Delta A_{min.} D/2.E_o.d$ where ΔA_{min} is the change of optical density per minute, D is the dilution factor, E_o is the molar extinction coefficient for NADH at 340nm. and d is the path length.

Baranowski and Niederland (1949) have given the value of the extinction coefficient as $E^{0.1\%}_{280nm} = 0.91$ and this allows the concentration to be calculated from an O.D. measurement on a suitably diluted solution. Using this, the specific activity is easily obtained.

Preparation of Rabbit Muscle Aldolase

Crystalline aldolase has been obtained by several groups: Taylor Green and Cori (1948), Baranowski (1939), Beisenhartz et al. (1953) and more recently, Czok (1960) and Eagles et al. (1969). Of these methods the most commonly used one is that of Taylor et al. (e.g. Taylor (1955, 1957)) possibly because it is the most straightforward. It was the Taylor method with modifications by Swenson and Boyer (1956) and Lai and Horecker (1967), Lai (1968) which was used in this study. Since the primary object was to obtain aldolase, the protein fractions containing only small amounts of aldolase were not kept.

pH measurements throughout this work were made using an EIL Model 23A direct reading pH meter fitted with a Jena Type 9259/81 micro-dual combination electrode. It was standardised before use with 0.025M KH_2PO_4-0.025M $Na_2HPO_4.2H_2O$ secondary standard solution pH 6.86 at 20°C which itself had been standardised with 0.05M potassium hydrogen phthalate primary pH standard solution whose pH is defined as 4.00 at 20°C (Dawson et al. (1959)). The salts used for these solutions were Analar grade and the potassium hydrogen phthalate and KH_2PO_4 solids were air dried at 100°C for 2-3 hours before use.

A laboratory rabbit was concussed and as much blood as could be bled from the corpse was removed. The skeletal muscle, mainly from the back and rear legs, was excised and stored at -10°C until required. About 0.6-1.0 kg. was usually obtained. A large quantity of .03M KOH/ 0.001M EDTA was cooled in the cold-room to about 3°C. A chilled Kenwood meat-mincer was then used to mince the muscle and it was found that, although attempts were carried out to render the mince to a more homogeneous consistency using the 'blender', further grinding was

(a) PURIFICATION - Colorimetric assay

A.S. Fraction	Total Activity (units)	Protein (g)	Specific Activity (units/mg)
Total	34,400	22.0	1.5
0 - .4 sat.	1,100*	3.8	0.3
.4 - .5 sat.	2,700	0.7	3.9
.5 - .54 sat.	23,220	2.3	10.6
.54 - 1.0 sat.	7,380	16.1	0.5

* By difference

(b) CRYSTALLISATION - Dynamic assay

	Specific Activity (units/mg.)	Grammes Protein
First crystals	4.6	2.3
After third recrystallisation	9.6	1.4
After Dialysis	11.5 ± 0.6	1.1 ± 0.2

TABLE III.2 Sample protocol for the purification and recrystallisation of rabbit muscle aldolase.

unnecessary. The meat was extracted twice with equal volumes (1ml. for 1g.) of the KOH solution by stirring the slurry for 15-20 minutes and then straining through cheesecloth (or centrifuging it for 30 minutes at 2000 r.p.m. in a M.S.E. Mistral 6L refrigerated centrifuge). This yields a total volume of about 200-250 mls. per 100g. of fresh muscle weight. The pH of this extract was adjusted to 7.5 with 0.1M KOH and then treated with a solution of ammonium sulphate (A.S.) saturated at room temperature and then titrated to pH 7.5 with dilute NH_4OH, to bring the salt concentration to 40%. The A.S. solution was made up by dissolving slightly more than four moles per litre of Analar crystals in hot distilled water and then filtering through a heated Buchner funnel to remove the dust and other insoluble dirt which seemed to be present in the commercial material supplied both by Hopkins and Williams and by British Drug Houses. On cooling, the pH was adjusted with concentrated ammonia solution, measurement being made on a tenfold diluted sample. The precipitate formed by the addition of A.S. was spun down and discarded. The saturation was taken to 50% with more A.S. solution (33 mls./100 mls. of original extract) and then set aside to precipitate. The solid was removed by centrifugation after about 30-40 mins. and the saturation increased to 54% (20mls./100mls. of solution). This solution was then left at 3°C to crystallise. Three to four days later a 'sheen' of aldolase crystals had become evident on swirling the flask. A sample protocol is shown in Table III. 2.

Having obtained crystalline aldolase, it was necessary to recrystallise it. This was done at pH 6.5 because Baranowski and Niederland (1949) state that better separation can be obtained at the lower pH. The recrystallisation procedure was as follows.

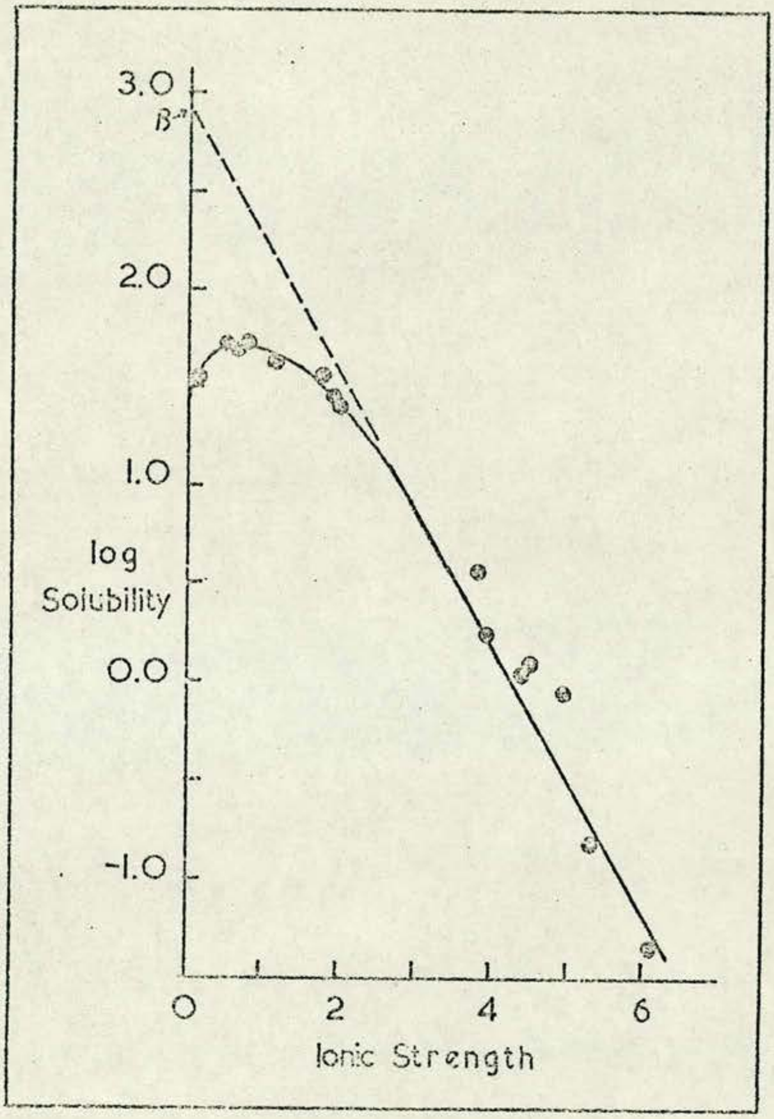
A solution containing about 10 mg./ml. was taken to the first signs of a permanent turbidity (about 45% saturation) and then left. This procedure was done at $3^{o}C$. Some twelve hours later the crystals were centrifuged and then resuspended in 2M A.S. to wash them. They were again spun down, the washing supernatant drawn off and the crystals then redissolved in a small quantity of twice distilled water, there being enough salt still associated with the precipitate to allow the enzyme crystals to dissolve. The solution was then filtered through Whatman's No. 1 and then reprecipitated as before. This procedure was repeated twice more. Taylor (1957) states that all contaminating enzymes are removed by several recrystallisations but Mehler (1965) has observed that overnight dialysis against cold distilled water, after recrystallisation is required to eliminate the effects of the major contaminant, triose phosphate isomerase. Dialysis was carried out as prescribed yielding aldolase with a specific activity of about 115 units/mg. which is higher than that of the available commercial preparations which are typically about 9-10 units/mg. However, it was noticed that other workers used recrystallised aldolase with an activity, usually, of between 15-20 units/mg. (Richards, Rutter and Woodfin (1961), Mehler (1963) for instance, although Lai (1968) whose method was employed here obtained typically 12-14 units/mg. Webb (1963) has said that the presence of salts, especially borates and phosphates, will usually inhibit competitively enzymes which act on phosphate-containing substrates. Also, the buffer may affect the ionisation of groups at or near the enzyme's active site altering the binding of the substrate, and hence the activity. Fig. III. 3 shows the data of Dounce, Barnett and Beyer (1950) showing the effect of different buffers on the activity of aldolase at various pH's. As can be seen, the presence of phosphate ion causes the activity to drop. Now, collidine, the buffer in which the assay is carried out, was present as the major constituent but phosphate was always used to dilute the enzyme solution before the activity was measured so that it, too, would be present and would already have been bound before addition. This, plus the large dilution necessary as mentioned on page 44 could explain the slightly lower values obtained in both procedures.

The aldolase obtained above, plus some from a similar preparation, was then used for the examination of the "crystallisability" of the protein, since, before any X-ray work can be done, single crystals of reasonable size (0.1-0.5mm) must be grown.

Factors Affecting Crystallisation

The way used for the growth of protein crystals depends on causing the protein to become insoluble in aqueous solution by one of two methods: either by changing the salt concentration or by altering the dielectric constant of the solution by the addition of organic solvents. A protein in dilute buffer solution can be considered to be a sphere with charges spread over its surface, each of which is 'solvated' by water molecules and/or salt ions, such that it remains in solution. If anything is changed in the solvent, the degree of solvation of the protein will alter causing the surface charges to redistribute themselves and this change may cause the protein to become insoluble. With a molecule as complicated as a protein, what actually has occurred when it becomes insoluble is not fully understood because there are so many variables.

The alteration of the dielectric constant of an aqueous solution can best be done by adding an organic solvent such as acetone or



<u>FIG. III.4</u> Relationship between log solubility and ionic strength for a typical protein. Data is from Green (1931, 1932) for carboxyhaemoglobin in $(NH_4)_2SO_4$ at pH 6.6 and 25°C. Solubility is in g./1, and ionic strength is in moles/1.

ethanol. The effectiveness of alcohol-water mixtures as a means of protein extraction was found by the eighteenth century chemists but they also discovered that, in the dissolution process, the protein became modified or denatured and this is the main drawback of the method. However, successful procedures have been developed by Cohn (1946), Askonas (1951) and others for fractionation using temperatures just above freezing, and low ionic strengths. Fractionation or crystallisation by this means allows exploitation of the state when the protein is most sensitive to electrolytes, since its solubility depends greatly on its permanent and induced dipoles which are usually highly specific to it. The other advantage is that the protein is not present in large concentrations of salt which may subsequently have to be removed if, for example, the preparation is required clinically. The drawback is the one already mentioned of the great tendency to denature if the temperature is allowed to rise even slightly. Although use of organic solvents can produce single crystals such as serum albumin (Cohn (1946)) and chymotrypsinogen (Kraut et al. (1962)), it was decided to concentrate on salting out for two reasons. First, a salting out procedure had already been used for the initial purification and was therefore more familiar, and second, a cold room was not readily available.

As already mentioned, the other and more common method is that of varying the salt concentration. The solubility curve of a typical globular protein is shown in Fig. III. 4. This shows that from a protein solution in a dilute buffer, the solubility can be decreased by either increasing or decreasing the ionic strength, all other variables being kept constant. Many proteins can be precipitated in salt free water and some, β-lactoglobulin, for example, will grow good single crystals (Aschaffenburg et al. (1965)). Aldolase will precipitate from salt free solution it was found both by Askonas and the author, but it was decided not to pursue this method in the first instance. This was because three forms had been reported by salting out (Baranowski (1939), Taylor et al. (1948), Beisenhertz et al. (1953), Wolf and Leuthardt (1957)) and none by salt free precipitation. Also, in salt free solution the pH tends to that of the isoelectric point of the protein which may not give a suitable form for X-ray work. Adding heavy metal ions could have a salting in effect on the crystals or, as discussed by Aschaffenburg et al., could change the crystal form. The method of growing such salt free crystals requires the slow removal of salt by dialysis which tends to cause them to adhere to the membrane, where they might even react with the thiols in the cellophane, making their removal without breakage, rather awkward. Finally, dilute and salt free protein solutions make ideal feeding grounds for bacteria unless suitable precautions are taken. Whilst these latter difficulties could be overcome, it was felt more hopeful to try salting out.

In growing crystals by salting out, there are four variables to be controlled: temperature, pH and the concentrations of salt and protein. Also, the nature of the salt is of importance. As can be seen from Fig. III. 4, at ionic strengths above 4 in this instance, the curve becomes linear and Cohn (1925) has expressed it mathematically as

$$\log s = \beta' - K'_s . \Gamma/2$$

where s is the concentration of protein in g./l., β' and K'_s are constants and $\Gamma/2$ is the ionic strength. This expression is not peculiar to proteins, being applicable also to most organic substances. K'_s is independent of pH and temperature but varies with

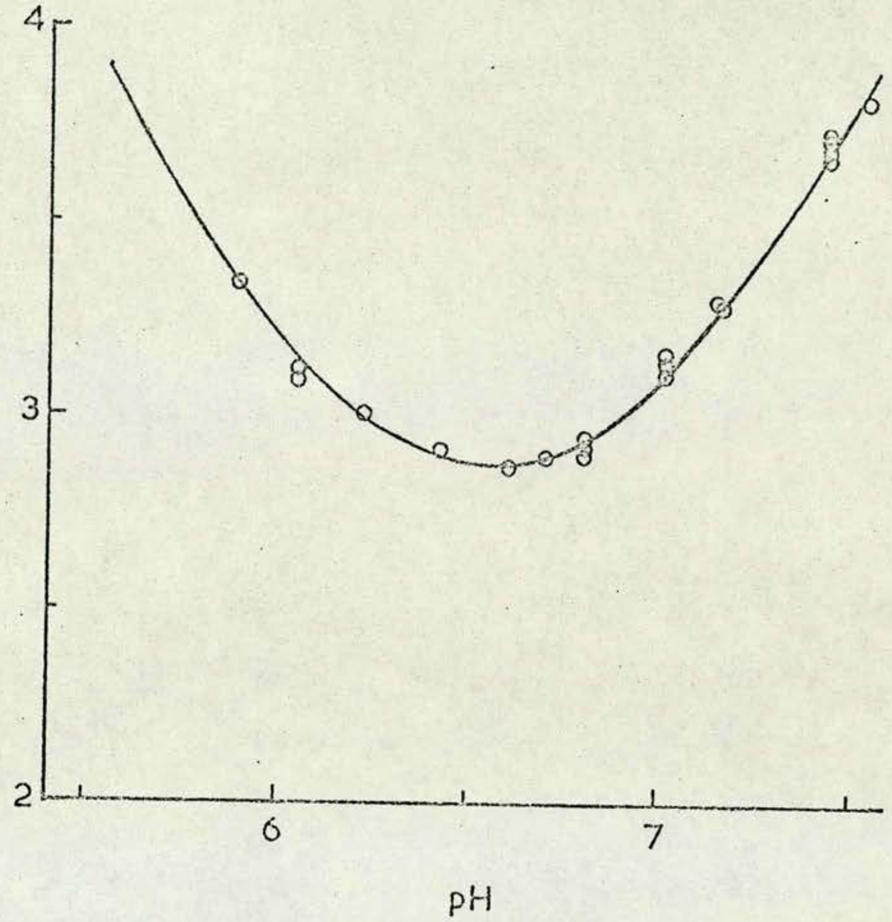


FIG. III.5 Effect of pH on solubility. The data are from Green (1931,1932) for carboxy haemoglobin at 25°C.

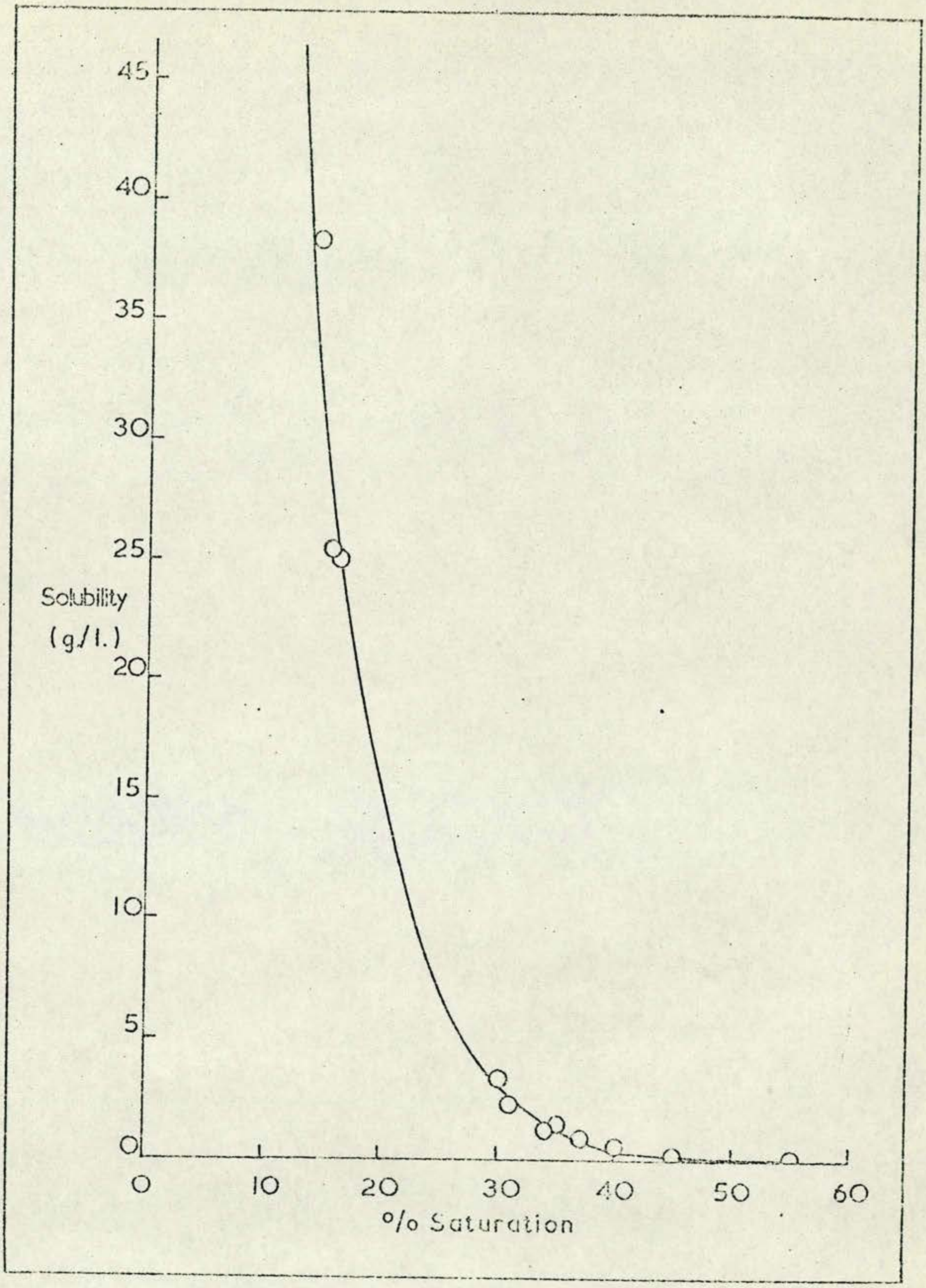


FIG. III.6 Carboxyhaemoglobin data from Fig. III.4 replotted as solubility (g./l.) versus percentage saturation of $(NH_4)_2SO_4$ at pH 6.6 and 25°C, illustrating the need for a knowledge of the protein concentration in salting out procedures.

the type of salt and the protein. β', the log of the solubility extrapolated to zero ionic strength, depends on pH, temperature and the protein but is almost independent of the salt. What is the effect on the solubility of each of these variables?

Fig. III. 5 shows the effect of varying the pH alone. The minimum, near pH 6.5 is close to the isoelectric point, a feature which is common to most proteins. However, some proteins have more than one minimum. Sorensen and Sorensen (1933) have shown two in the case of carboxyhaemoglobin: the one shown at pH 6.6(Fig. III. 5) and one at pH 5.5. Enolase has three (Bealing et al. 1960)). Because the change of solubility with pH is small close to such minima, crystal growing at such a point will require less stringent control and, as in enolase where more than one minimum exists, each may correspond to a different form.

Many proteins have negative heats of solution so that crystallisation can be induced by allowing their solutions to warm up slowly. Taylor et al. (1948) found that aldolase crystallised as plates when allowed to warm from $3^{o}C$ to room temperature. This effect can also be used to great advantage in protein fractionation, since two proteins with the same solubility at $0^{o}C$ will most likely have differing solubilities as the temperature is raised.

If the data in Fig. III. 4 are plotted as shown in Fig. III. 6, it can be seen that the point of precipitation of the protein depends on its concentration. This is of great practical importance in fractionation, especially, since it means that a protein may not be present within the limits expected. Thus, in following a reported separation procedure, the concentration of each species required, must be monitored and the appropriate modifications made. This also explains why the percentage saturation required to recrystallise aldolase is lower than that of the fraction in which it first separated.

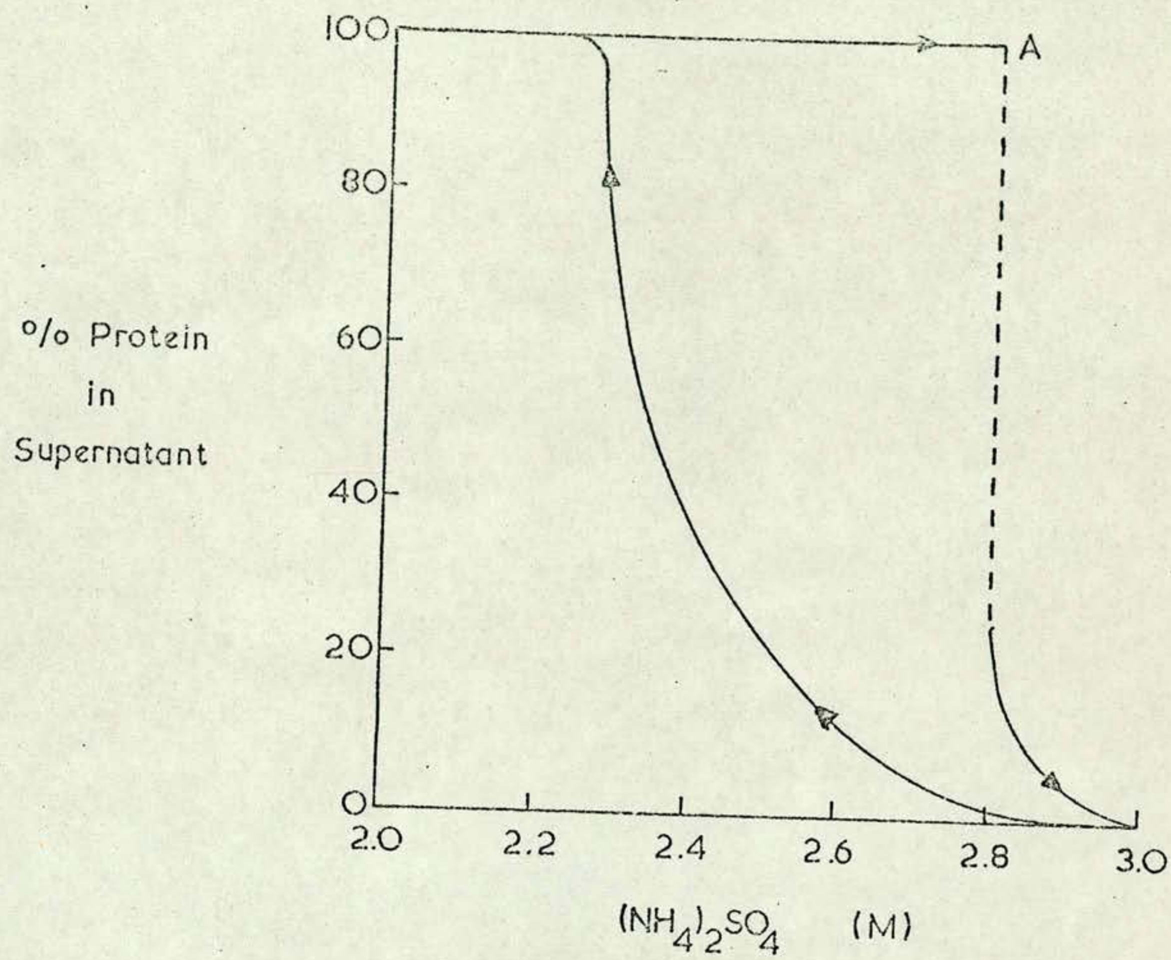


FIG. III.7 Enolase precipitation at pH 7.8, 0.02M triethanolamine-HCl, 1.2mM EDTA, rate of addition of A.S. was 0.05M/hr at 3°C. There was a pause of 24hrs at point A. After Bealing et al. (1960).

	Potassium phosphate	Sodium sulphate	Ammonium sulphate	Sodium citrate	Magnesium sulphate
β'	3.01	2.53	3.09	2.60	3.23
K'	1.00	0.76	0.71	0.69	0.33

TABLE III.3 Values of β' and K' for carboxyhaemoglobin for various salts. Data from Dixon and Webb (1961).

What about the salt used for salting out? Hofmeister (1887) has postulated that the ability of a salt to precipitate a protein is related to the charges on both cation and anion such that increasing cationic and decreasing anionic charges become less effective in precipitating proteins. This is also borne out to some extent by the values K_s' for carboxyhaemoglobin in different salts as shown in Table III.3. The data are from Dixon and Webb (1961), the values of β' being included for comparison. It can be seen, then, that the precipitants are usually salts with polyvalent anions and monovalent cations. Of the ones shown, A.S. has probably been the most used since it is readily available in a pure form (though recrystallisation is a wise precaution) and it has a high solubility ($\sim$3.9M at 20^{o}C). One major drawback is that it has little or no buffering action so that control of pH requires small concentrations of buffer. Phosphates do have this buffering property but here the problem is the differing solubilities of the mono- and di-hydrogen salts of potassium and sodium, those most commonly used.

Bealing et al. (1960) have illustrated another facet of crystallisation by salting out. Fig. III. 7 shows the effect of adding solid A.S. at a constant rate to a solution of enolase, and then the dissolution of the crystalline precipitate formed. At the point, A, the solution is supersaturated but it is not until this point that any material appears in the form of a faint amorphous precipitate. On leaving for 24 hours, some 80% of the protein has appeared as crystals. The hysteresis exhibits a fact, common to most proteins, that an amorphous precipitate is considerably more soluble than a crystalline one and it is possible that the fine amorphous material first appearing acts as the supply of nuclei

pH	$-[\alpha]_D^{20}$
5.50	15.7
6.05	14.9
6.50	16.1
7.40	14.6
7.95	15.6
8.60	8.8 *

* Solid material present ; probable denaturation.

TABLE III.4 The effect of pH on the specific rotation ~~activity~~ of rabbit muscle aldolase in 0.1 M phosphate buffer.

Note: The effect of ion binding can be seen if the above values are compared with those of Jirgensons (1961) in 0.02 M phosphate.

pH 6.6 $-[\alpha]_D^{20} = 23.7$

pH 6.1 $-[\alpha]_D^{20} = 12.6$ (+ 0.0 M Sulphate) on which the crystals will grow. Thus, in recrystallising a protein, it is necessary to go on adding salt until the first cloudiness appears and then to leave the solution for the crystals to grow.

Crystallisation of Aldolase for X-Ray Work.

(a) Preliminaries

From the published work on the enzyme, it was known that two, and possibly three, crystal forms existed: hexagonal bipyramids at pH 6.5, plates at pH 7.5 and needles as the precursor of both these forms. Tanford, Bunville and Nozaki (1959) found that there is a marked conformational change in β-lactoglobulin as the pH is raised from 6.5 to 8.0 and Aschaffenburg, Green and Simmons (1965) have related this to the three crystalline forms obtained in salt solution. As a preliminary, therefore, it was considered worthwhile examining the optical rotation as a function of pH to see if a similar conformational change might be responsible for the different forms and, if so, determine the pH of the change.

10mls. of a suspension of aldolase were dialysed against 3 l. of 0.1M NaH_2PO_4/NaOH buffer, pH 5.5 overnight. The concentration of protein, its pH and the optical rotation were measured, the latter being measured with a Perkin-Elmer 141 Polarimeter in a 10cm. cell. The protein was then returned to dialyse for a further day against the same solution which had had its pH raised by 0.5 pH units. The process was repeated up to pH 8.6. The results are shown in Table III. 4 from which it was obvious that no dramatic change, similar to that in β-lactoglobulin, was present so that no such indication of the limits of production of the two crystal forms could be obtained.

It was found that, as for enolase, that the precipitation point for the amorphous material was at a higher salt concentration than that of the crystalline. That is, if a solution is taken to the first signs of cloudiness and left, a far heavier, crystalline precipitate appears. This was determined by dialysis of a small amount of protein in a well-stirred solution of A.S. to which solid could be added to increase the concentration without disturbing the aldolase. The difference between precipitation and redissolution was found to be about 0.1M, a value which appeared to be independent of protein concentration and pH. It is just at the initial precipitation point that the protein should be kept for the growth of large crystals since the solution is supersaturated with respect to the crystalline form and on the point of saturation to the amorphous. Thus, the growth of protein crystals by salting out can be considered as being a slow change from the rapidly formed amorphous state to the more stable crystalline one and this will require sufficient salt concentration to produce solid which can crystallise on nuclei (which may or may not be the same material: dust or unevennesses on the glass walls will suffice) but yet not too strong as to forbid redissolution. The process can be summarised as shown:

$$\text{Solute} \underset{}{\overset{\text{Rapid}}{\rightleftharpoons}} \text{Precipitate} \overset{\text{Slow}}{\rightleftharpoons} \text{Crystals}$$

For aldolase, Wolf and Leuthardt (1957) maintain that there is an extra stage at low temperatures with an equilibrium between two crystal forms. This is shown below:

$$\text{Solute} \overset{\text{Rapid}}{\rightleftharpoons} \text{Precipitate} \overset{\text{Slow}}{\rightleftharpoons} \text{Needles} \overset{\text{Slow, temp.?}}{\rightleftharpoons} \begin{array}{l}\text{Hexagonal}\\ \text{bipyramids}\end{array}$$

It is because the optimum conditions for this process are fairly precise, that it is necessary to set up a series of crystallisations varying one parameter, usually salt concentration, by small increments.

Accordingly, each crystallisation was made up as a series of eight in 4 ml. glass vials with polyethylene stoppers. The vials were all cleaned with chromic acid, rinsed with copious quantities of distilled water and dried in an air oven at 110°C. The procedure employed in making up each of these tubes was as follows:

(a) Pipette into each tube the required volume of salt adjusted to the appropriate pH.

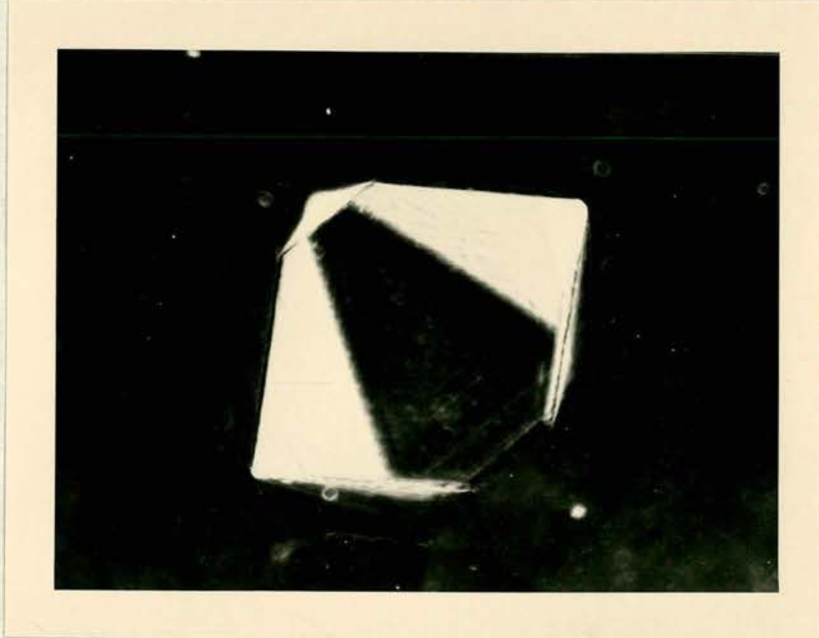
(b) Pipette into each tube the required volume of water and mix.

(c) Carefully add the protein drop by drop onto the centre of the surface, not down the sides of the tube, and mix.

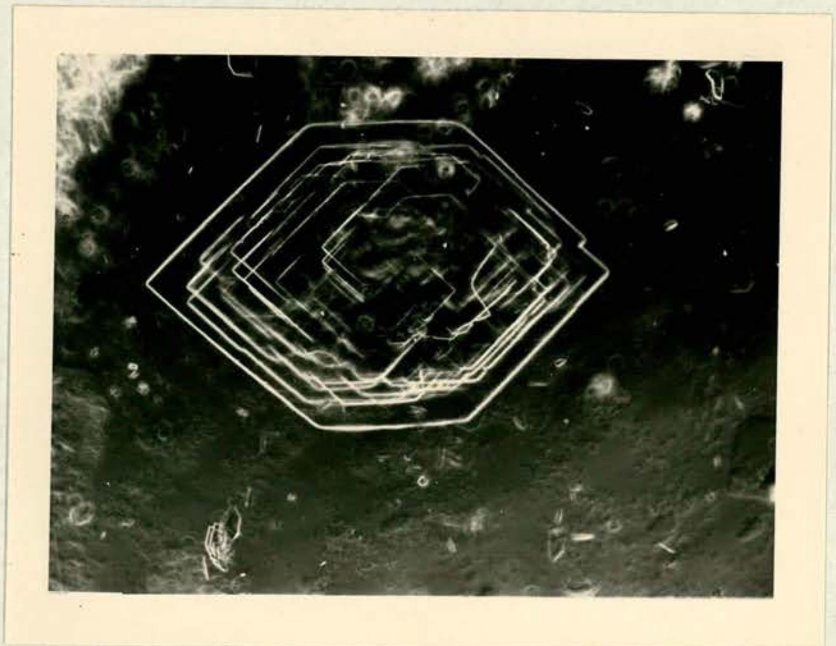
(d) Check the pH and adjust if necessary.

Step (d) poses a problem. The protein has a certain buffering action of its own sufficient to cause an appreciable pH variation, even in strong phosphate buffer, but which usually takes a day or so to be noticed. This is probably the slow solvation of the protein by the 2M salt solution. In strong salt solution any pH measured must be inaccurate because glass electrodes are only accurate if used in dilute solutions of less than 0.5M. Because the volume of the tubes was small, the material precious and the glass electrode only semi-micro, dilution of the final mixture was considered best avoided. The adjustment of the salt solution beforehand was done on a ten times diluted sample and, after making up the tube, the pH measured directly was usually about 0.5 pH units different. Once the pH of the tube had settled, the adjustment where necessary could be made using this difference. This was not, however, wholly satisfactory but the best that could be achieved in the circumstances.

It was found that, in A.S. solution at pH's 6.0, 6.5 and 7.5,



<u>FIG. III.8</u> Photomicrograph (x75) of the low pH form of aldolase showing the vestigial [0001] faces.



<u>FIG. III.9</u> Photomicrograph (x75) of the high pH form of aldolase.

although microcrystals formed as shown by the "sheen" on inverting the tubes, no large crystals formed even after two or three months. In fact, the aldolase formed a thixotropic gel both in the tubes and in the stock suspension even at protein concentrations of 1-2mg/ml. This effect has also been noticed by Disteche (1948). This intermolecular aggregation could effectively stop the rearrangement or solution-precipitation required for larger crystals to grow. It was also considered likely that the lack of buffering of the A.S. would create unstable enough conditions to stop growth. However, it was found that the use of 0.1M phosphate buffer did not improve matters and large crystals did not appear in this medium either. Before trying out the effect of altering the buffers with A.S., it was decided to try using phosphate alone since this was known to produce crystals of haemoglobin (Perutz (1968)), β-lactoglobulin (Dr. Green, personal communication). Solutions of 4M NaH_2PO_4-K_2HPO_4 at pH's 6.0, 6.5 and 7.0 were made up and these used to set up series of tubes ranging from 1.70M to 2.30M in salt and containing 1.5-3.0mg./ml. of aldolase. These yielded hexagonal bipyramidal crystals after three weeks at 3°C. It was found that the largest crystals appeared between 1.85M and 2.15M and that in the tubes containing these, there was little or no visible amorphous or microcrystalline material. The crystals varied between 0.1 and 0.6mm. in length and are shown in Fig. III. 8.

Having achieved a moderate amount of success in phosphate, it was decided to try to grow the high pH form from the same medium. Using a molarity of 1.95 in phosphate and a protein concentration of 2.8mg./ml., the pH was varied between 7.2 and 8.2 with 4M NaOH and a series of tubes set up at 3°C. After six weeks, crystals had appeared in some of the tubes including the malformed plates shown in Fig. III. 9. Surprisingly, the tubes at higher pH had hexagonal bipyramids as did those just below. The pH was measured directly as 7.4, and attempts to repeat the conditions only once yielded plates which, like those obtained before, were thin and malformed. The pH was measured at 7.48 and the molarity at 2.05. Usually bipyramids were formed and at pH values above 7.6 no crystals were obtained. Could it be that the range of plate crystals in phosphate was as narrow as 0.2 pH unit all other variables being kept constant?

A further examination of this aspect was about to be undertaken when it was discovered that a group in Bristol had produced crystals of a p-chlorobenzenesulphonic acid derivative at pH 7.4 from ammonium sulphate and that these had been sent to Oxford where work was being started. It was decided then to terminate the project since, as will be seen in the following chapter, structure determination on the low pH form would be a venture of unlikely success.

Nevertheless, a series of possible lines of research had been considered before the other group's work became known. It was thought possible that crystals should be able to be produced from A.S. because of the reports of Taylor et al. (1948) and it was realised that the buffer was the most likely factor to be the cause of the trouble. Phosphate, it was found, could bind to the molecule and this could affect the packing in such a way as to render the plate-like form unstable. But there are other buffers which were going to be tried, namely 'tris', acetate and citrate. Not only buffers could have been changed but also salting-out materials were considered worth trying such as the sulphates of lithium, sodium and magnesium and also ammonium phosphate. Another possibility was the controlled reduction of salt concentration in an attempt to grow salt free crystals.

Had these possibilities produced nothing, it was thought that some minor modification of the protein would cause it to crystallise. For instance, the substrate can be reduced with borohydride becoming attached to a lysine residue in the active site (Lai et al. (1967)). Carboxymethylation of the various cysteine residues by the method of Crestfield, Moore and Stein (1963) was another possibility and yet another was treatment with carboxypeptidase which removes the four terminal tyrosines from the carboxy termini (Morse et al. (1967)). Would these possibilities produce crystal forms which could be used for structure determination?

CHAPTER 4

ALDOLASE: X-ray work

Physical and Optical Properties of Aldolase Crystals

The larger crystals of aldolase obtained as described, were examined under the microscope. Typically they were 0.1-0.6mm. long, exhibited moderate birefringence and had faces of the form $\{hh\overline{2h}l\}$ or $\{h0\overline{h}l\}$. Some of the crystals had vestigial faces of the form $\{0001\}$, these occurring in tubes of pH 7.1. Careful measurement of the angles of the face opposite that on which the crystal was resting, showed that it was almost horizontal and therefore showed the angles between the edges of the crystal were close to their true values. These angles allowed the axial ratios for both possible sets of axes to be calculated, as shown in Fig. IV·1 (a). The final assignment of the axes could not be made until the unit cell parameters were known. This, then, was the next stage.

Mounting the Crystals for X-Ray Work

Protein crystals have to be kept moist or else, if allowed to dry, the salt precipitating internally so disrupts the structure that no diffraction pattern is obtained. This is done by allowing the crystal to adhere to the wall of a fine glass or quartz tube of diameter 1-2mm. by the capillary action of the mother liquor (Boyes-Watson, Davidson, Perutz (1947)). Almost all of the mother liquor can be drawn off with fine pieces of cotton, which can also be used to move the crystal to any required orientation. To maintain a moist atmosphere round the crystal, a piece of cotton soaked in mother liquor is placed at each end of the capillary and finally the tube is sealed with small blobs of wax. In this way, the crystal can be kept moist and it is relatively free from material which will spoil

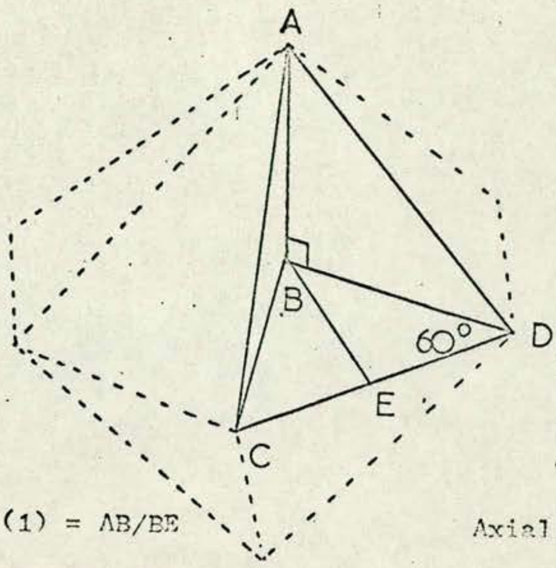


FIG. IV.1a

Figure ABCD is 1/12 of the complete hexagonal bipyramid shown dashed. B is its centre.

Axial ratio (1) = AB/BE

Axial ratio (2) = AB/BC

BDC is an equilateral triangle such that BD = 1 unit.

Then, ratio (1) = tan(cos-1(3/tan(ACD)))

and ratio (2) = ratio(1)*sin 60°

Crystal measurements gave

CAD = 43.7° : ACD = 68.4° ; ADC = 69.5°

Mean ACD = 69.0 ± 1.0° giving

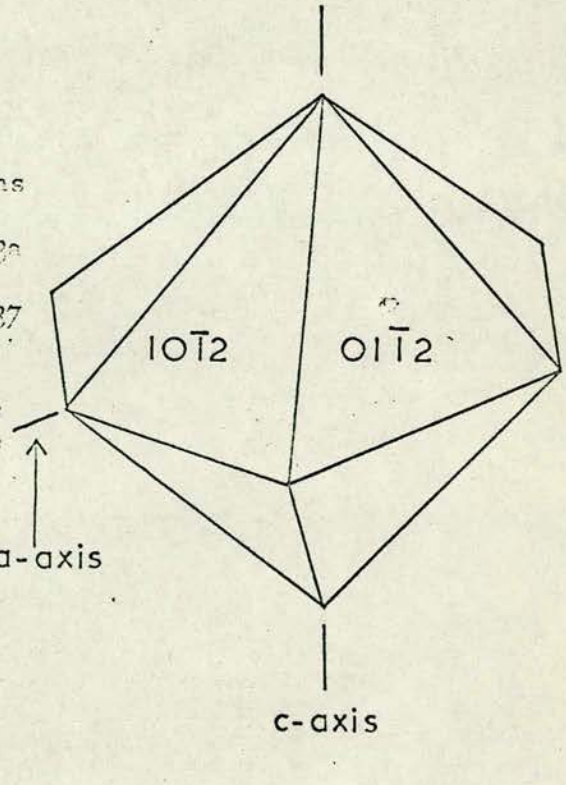
<u>Ratio (1) = 1.13</u>

<u>Ratio (2) = 0.98</u>

FIG. IV.1b From the cell dimensions

c/a	c/2a	c/3a	2c/3a
2.05	<u>1.02</u>	0.68	1.37

This allows the axes to be assigned as shown, bearing in mind that the a-axis is parallel to a crystal edge.



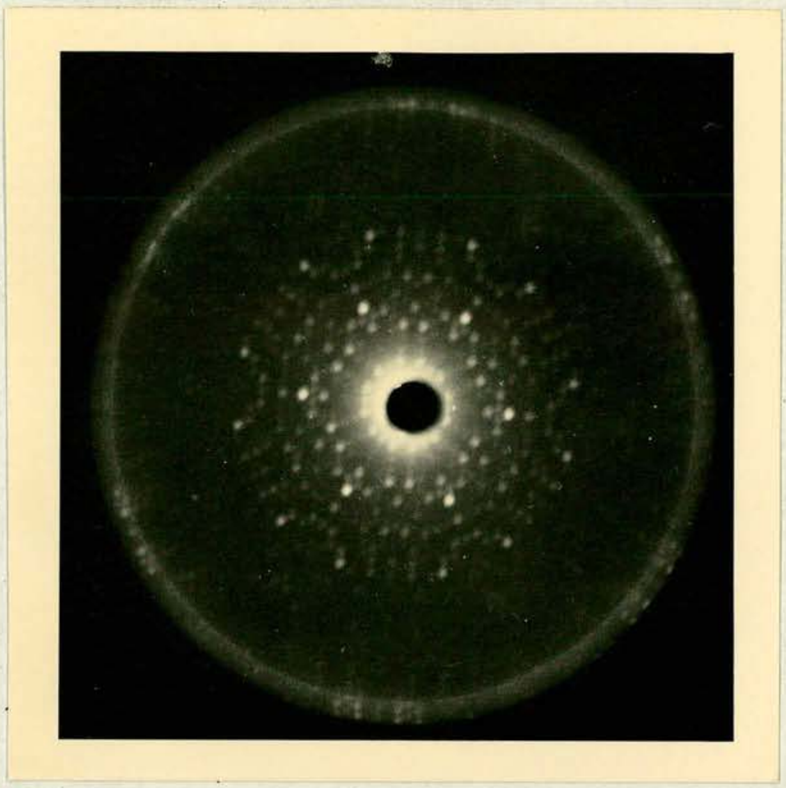


FIG. IV.2 [0001] precession photograph of rabbit muscle aldolase, pH=6.6.

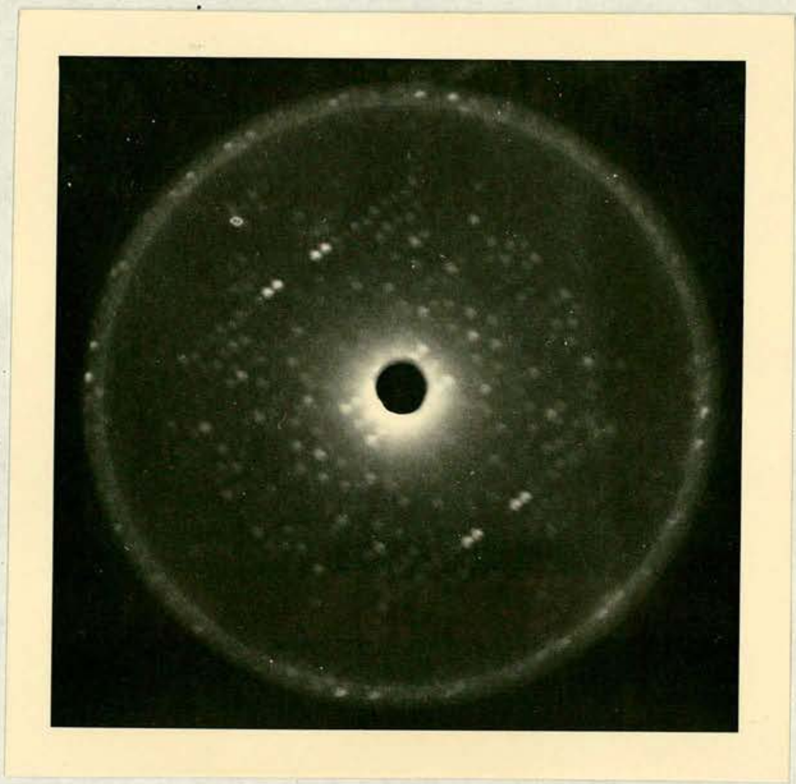


FIG. IV.3 [$01\bar{1}0$] precession photograph of rabbit muscle aldolase, pH=6.6.

the diffraction pattern, whilst still allowing it to be mounted on a normal goniometer head with plasticine in the usual way.

Determination of Cell Dimensions.

Photographs were taken on Ilford Industrial G film using copper radiation, a Supper precession camera with f=10cms, a collimator of 0.2mm. and a Phillips fine focus tube on a Phillips PW1010 generator. This was run at 40kv and 18mA. Setting photographs were taken using unfiltered radiation.

It was found that the crystals were very poor diffractors of X-rays. This made setting very time consuming, since a Laue photograph took 30 minutes, and a small angle precession photograph 60-90 minutes. A full 8^{o} precession photograph, of rather poor intensity, took some 20 hrs. Also, the life of the crystal was found to be about 20-24 hrs. meaning that one crystal could give one photograph if setting could be achieved in two or three attempts. The use of Polaroid film reduced the exposure by a factor of three and careful preliminary setting by eye using a microscope usually enabled one photograph to be taken per crystal.

Crystals at pH 7.1 were found to have a shorter life than those at lower values, again indicative of an increasing instability. However, within the accuracy of measurement of the films, no change in cell dimension over the range pH 6.0-pH 7.1 was found.

Photographs were taken down the apparent sixfold axis (Fig. IV. 2) and, perpendicular to this parallel to an edge, (Fig. IV. 3). Because Fig. IV. 3 had spots almost overlapping, i.e. a very long axis and because it was thought possible that more than one layer was being allowed through the screen, still photographs were taken without a screen both down the [0001] axis and the [01$\bar{1}$0] axis to clarify this

	a (Å)	c (Å)	V (Å^3 x 10^{-6})
Laue photograph	162.8 ±0.5	335.4 ±1.3	7.69 ±0.07
Precession photograph	163.5 ±0.6	335.0 ±3.0	7.76 ±0.15
Mean	163.2 ±0.6	335.2 ±2.1	7.73 ±0.11

<u>TABLE IV.1</u> Unit cell dimensions of rabbit muscle aldolase, pH 6.0.

point. Measurement of the films was done with a travelling microscope. This enabled the cell dimensions to be calculated from both types of film. From the relationship

$$n\lambda/d = (1-\cos\Theta)$$

where d is the axial spacing, n is the "ring" number, λ the wavelength and Θ the angle whose tangent is the ratio of the "ring" radius to the focal length of the camera, it can be seen that a plot of n versus $(1-\cos\Theta)$ should give a straight line of slope λ/d. The value obtained was found by a least-squares fit of n and $(1-\cos\Theta)$. The results fom these methods are shown in Table IV.1. As a further check of the value obtained for the c-axis, the rows of constant k were micro-densitometered and the peaks superimposed on a grid of spacing corresponding to a spacing of 335Å. It was found that most the peaks did fall on the grid, those that did not being put down to poor spot shapes causing the maximum not to appear at the centre. Using this value of the spacing up the c-axis, calculation showed that no upper level spots were being recorded on the zero level photograph, the annulus having a width of 2mm. Thus, the value of the c-axis was taken as 335Å.

From these values, it was possible to calculate the axial ratio and to identify the axes and the faces of the crystal. The faces turned out to be of the form $\{10\bar{1}2\}$. (Fig. IV. 1 (b)).

Determination of the Space Group

By virtue of a complete diffraction pattern's always having a centre of symmetry (neglecting anomalous scattering), a threefold axis will appear as a sixfold one on a zero level photograph. To resolve this ambiguity an upper level photograph was taken. (In fact, layer 5 was taken since, from the Laue, it looked most strong.) This showed there to be a sixfold axis. Similarly, an upper level photograph

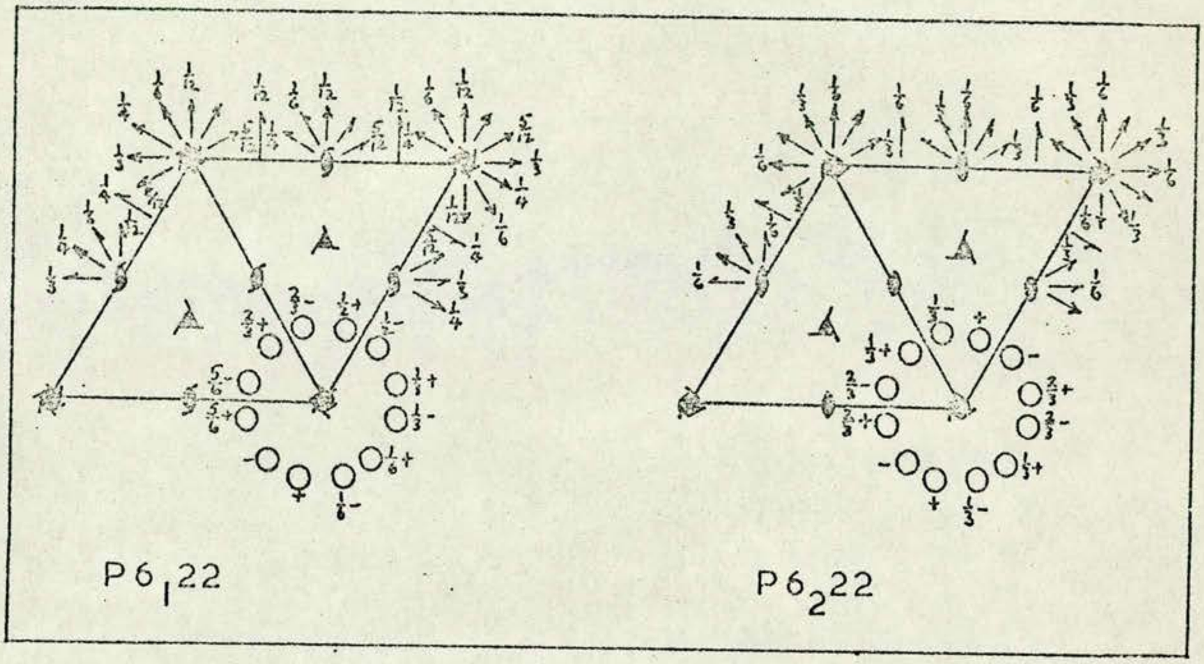


FIG. IV.4

of an [$01\bar{1}0$] zone showed there to be two twofold axes at right angles.

All space groups which have centres of symmetry, mirror- and glide-panes are forbidden to protein crystals because protein molecules are made up of homoenantimorphous amino acid residues. Thus, from International Tables, Vol. I, the only space groups possible having a sixfold axis perpendicular to two twofold ones, are $P6_n22$ with n=1,2,3,4,5.

Considering the great length of the unit cell, this type of axis is favoured rather than a pure rotation one. From the photograph showing the 0001 row, only those spots appear which satisfy l=6n. Careful examination revealed that no spots <u>not</u> satisfying this condition appear, although if they were very weak they could escape detection owing to the quality of the film. This, then, pointed to a 6_1 (6_5) screw axis.

Thus, the space group would appear to be $P6_122$ ($P6_522$), at least to 5.5Å resolution (Fig. IV. 4).

Determination of the Number of Molecules and Discussion.

The determination of the number of molecules in the unit cell requires the density, the water content and the salt content of the crystals to be found. The measurement of each of these quantities will be dealt with in turn.

(a) The Density

The method employed here is that using a gradient column as described by Low and Richards (1952). A preliminary determination using a toluene/CCl_4 column calibrated with drops of liquid of varying S.G. gave the protein density as about 1.3g./ml. A more precise measurement was made with a xylene/bromobenzene gradient made in the following manner. Sufficient xylene and bromobenzene were mixed together to give a homogeneous solution of S.G.=1.2. A glass cylinder was then half filled with bromobenzene onto which was carefully poured an equal measure of the less dense solution. The two layers were swirled together with a glass rod and the column left for several hours to equilibrate. A glass ball of diameter about 3 mm. was fixed to a length of nichrome resistance wire of diameter 0.001" and this then used to calibrate the column by weighing it at various heights. A graph was drawn which could be used to read off the density of a crystal floating anywhere in the column. Crystals were blotted free of mother liquor on a filter paper as well as could be in the shortest time, since they lose water to the atmosphere on standing. They were then placed onto the surface of the column with tweezers and a small bit of filter paper and pushed below the surface with the tip of a fine glass rod. The measurement of the height of the crystal was made by using a travelling microscope about 5 minutes after its introduction to the column. The upper and lower extremities were measured and the mean taken as the depth from which the density was determined. The mean value of the density of the crystals was 1.312 ± 0.006g./ml. The density of the mother liquor was found to be 1.309g/ml. which agreed with the observation that the crystals tended to sink.

(b) The Water Content.

This requires that separate crystals are weighed and dried. Because the crystals were small, three or four of the largest were used for each determination. Weighing was carried out on a Cahn Electrobalance, capable of weighing to .5μg., calibrated with a 5mg. weight. Small aluminium discs were made into boats for holding the crystals and dried in an air-oven to constant weight. Typically the weight of a boat was 3mg. and that of the crystals 0.4mg. The crystals were carefully blotted free of mother liquor and weighed every 30 secs. over a period of 5 mins. in order to extrapolate to zero time.

The seven determinations were dried in vacuo, initially at room temperature for 48 hrs., and then at 86°C for 72 hrs. over P_2O_5, by which time they had reached constant weight. The dry weights likewise were timed and extrapolated to zero time. The mean value of the water content which was removed was 42.5±1.9% of the original weight of the crystal plus water.

(c) The Salt Content.

The percentage weight of the salt in the mother liquor was determined by drying to constant weight at 110°C in an air oven, five 1.0ml. aliquots of a solution made up in the same ratio and of the same 4M K_2HPO_4 and NaH_2PO solutions which were used in the crystallisations. This gave a value of 31.5±0.7% of the weight of salt plus water.

Perutz (1946) has discussed the nature of the liquid occupying the space between protein molecules with particular reference to horse methaemoglobin. Within the crystal, the liquid can be considered as comprising two types: "free liquid" and "bound water". In other words, some of the water molecules are bound to the protein surface and are unaffected by the salt in the external solution but they can be driven off by heating them. The free liquid, on the other hand, is the salt plus water which, it is assumed, can percolate freely through the lattice and will therefore have the same salt content as the mother liquor. Perutz has determined the amount of bound water in methaemoglobin as 30% of the weight of the protein. As this value was calculated from an expression involving the partial specific volume, $\bar{v}_p$, and since most proteins have values of $\bar{v}_p$ of around 0.74ml./g. (Matthews (1968)) it was assumed that the amount of water bound to aldolase was 30% of the weight of the protein. This was most likely to be an overestimate because, although the molecular weight of aldolase about twice that of haemoglobin, the surface area does not double even if the subunits in aldolase are separate (see below), making the amount of bound water per gramme less, if it is assumed that the number of water binding sites per unit surface area is the same.

Approximate calculation of surface areas of Haemoglobin and Aldolase

Haemo: Surface area of sphere of radius 30Å = 11,300Å^2
(North (1959))

Aldolase: Surface area of four spheres of radius 20Å = 20,100Å^2
Penhoet et al. (1967) (maximum value)

Haurowitz (1950) showed that air-dried proteins retain some 4-10 g. of water per 100 g. of protein and North (1959) has suggested that the average of this value should be subtracted from the apparent weight of protein found in the unit cell. However, since the aldolase crystals were dried *in vacuo* to constant weight, it was thought that all of the water would have been removed. The number of molecules in the unit cell can now be calculated.

Density of crystals	$= 1.312 \pm 0.006$ g./cm.^{3}
Volume of the unit cell	$= 7.73 \pm 0.11 \times 10^{-18}$ cm.^{3}
Mass of the unit cell	$= 10.14 \pm 0.02 \times 10^{-18}$ g.
Mass loss on drying	$= 42.5 \pm 1.9\%$ of crystal mass
Mass of salt in mother liquor	$= 31.5 \pm 0.7\%$ of mass of salt and water
Bound water (assumed)	= 30% of mass of protein

Let p be the mass of protein in the unit cell.

Then 0.3p is the mass of bound water

The mass of mother liquor = Mass of unit cell - 1.3p

$= (10.14 \times 10^{-18} - 1.3p)$g.

Mass of water removed on drying = Mass of bound water + mass of water in mother liquor

$= 0.3p + 0.685\,(10.14 \times 10^{-18} - 1.3p)$

Solving this equation for p gives

Mass of protein in unit cell $= 4.465 \times 10^{-18}$g.

	Oxford	Edinburgh
Crystallisation medium	A.S., triethanolamine buffer, pH 6.0	Phosphate, pH 6.0
Unit cell dimensions	a=161.0Å , c=169.0Å , V=21.4x10^{5}Å^3	a=163.5Å , c=335.0Å , V=77.3x10^{5}Å^3
Space group	$P6_222$	$P6_122$
Density	1.250±0.006 g/ml	1.311 g/ml
Protein content	48.5 of total weight	43.8 of total weight
Molecules/unit cell	9	18
Molecules/asymmetric unit	3/4	3/2
Apparent molecular weight	130,000	148,500
V_m (using above M.W.)	2.75Å^3/dalton	2.80Å^3/dalton
V_m (using 158,000)	2.03Å^3/dalton	2.72Å^3/dalton

TABLE IV.2

These data are confirmed by Gorynnov, Andreyeva and Shpitsberg, Biofizika 14 1116 (1969)

There are 12 general positions in $P6_122$.

Thus, the weight of one asymmetric unit = 224,000 daltons
This gave a value of 1.42 molecules per asymmetric unit or 17 per unit cell, using the value of 158,000 for the molecular weight (see Table II. 2). However, the space group required there to be a multiple of six molecules so that the actual number appeared to be 18.

An alternative way of calculating the apparent molecular weight was suggested by a paper by Green, North and Aschaffenburg (1956). It was assumed that all of the water was free liquid and was therefore available to salt.

A similar calculation to the above was performed but this time x=0. Thus, the total percentage by weight of non-proteinous material in the crystal = 42.5.100/68.5 = 62.1%. This gives a weight of one asymmetric unit of 193,000 ± 4,000 daltons corresponding to 1.25 molecules or 15 per unit cell, a value equidistant between the two nearest allowed values of 12 and 18. This value, however, was taken as the lower limit since it was considered unlikely that no water would be bound (Bishop and Richards (1968)).

The indication, then, was that there were 18 molecules in the unit cell which would make the molecular weight lie in the range 129,000-149,000, a range much lower than the commonly accepted value of 158,000.

Matthews (1968) has defined the quantity V_m as the crystal volume per unit of protein molecular weight. Further, he goes on to examine most of the globular proteins on which X-ray data exist (to the time of his writing) and finds that the V_m are distributed with a positive skew having a median of 2.61 Å^3/dalton. He also points out that the lower and upper limits appear to be 1.68 Å^3/dalton and 3.53 $\AA^3$/dalton respectively, although one or two forms of some proteins do have values outside these. The values for 12, 18 and 24 molecules per unit cell of aldolase are shown below.

n/unit cell	V_m ($\AA^3$/dalton)
12	3.82
18	2.73
24	2.04

This pointed to either 18 or 24 molecules, but since the previous values were between 12 and 18, this added more weight to the earlier assignment of 18. This was taken only as a rough guide and could not have been used alone as evidence for the number of molecules. It must be remembered, however, that there was a large uncertainty on account of the assumptions made concerning the salt content so that as many independent indications of the value as possible were required.

The indications, then, were that there were 18 molecules per unit cell. This result was interesting in itself because it implied that at least one of the molecules was situated on a crystallographic dyad, or, put another way, there must be a molecular dyad parallel to (if not actually coincident with) the crystallographic one. In other words there were one and a half molecules per asymmetric unit. This meant that the molecule had to contain four subunits in accordance with the findings of Kawahara and Tanford (1966), Penhoet et al. (1967) and others unless the molecule had subunits which had molecular dyads themselves, a most unlikely feature.

From the point of view of further structure determination with this form of aldolase by use of the conventional techniques available to date, further work was considered likely to be fruitless. The length of the c-axis was such that collection of a three dimensional set of data would have been not only tedious but also difficult, the task being made no easier by the short life time of the crystals. Besides, location of the heavy atom sites by the normal vector methods would be hazardous. For instance, one site per molecule would give rise to 18 x 17 peaks in a three dimensional Patterson map so the task of unscrambling such a map would be almost impossible. If, as would be more likely, there were four sites per molecule, there would be 72 x 71 peaks to be interpreted. It was also considered pointless to try to determine the positions of the molecules in the unit cell by rotation-translation methods in projection for the following reasons. The [0001] photograph was taken down such a long axis with so many molecules in the cell that a projection would have been uninterpretable. The a-axis projection would also have been poor since the data from the photograph would be innaccurate, especially since the rows of spots nearly overlapped.

The only thing to be done was to abandon work on this form and to attempt to grow suitable crystals of the other form which, it was hoped would be more amenable to X-ray work. Indeed, work was in progress to this end when it was learned some months later of the investigation just being undertaken in Oxford.

It is of interest, however, to compare the X-ray work on the hexagonal form obtained by the group in Oxford, (Eagles et al. (1969), since there is a remarkable difference between their data and those reported here. A summary of their observations appears in Table IV.2 along with those of this work, for comparison.

Perhaps the most striking feature is the 6_2 screw axis in sulphate corresponding to a 6_1 screw axis in phosphate. That is, the translation for a turn of 60^o of helix is the same in both cases because the 6_1 axis operates over twice the distance. In sulphate, all of the molecules lie on dyads but in phosphate, with three times as many molecules, twelve of them are displaced from the dyad. This raised the question: were the cells related in some simple way? The protein volume in the larger unit cell is 3.1 times that of the smaller one, a further indication that the assignment made earlier is most probably correct.

Because the molecule comprised four subunits, the most likely arrangement was thought to be tetrahedral rather than square planar since extra stability would be gained if each subunit was in contact with all others rather than only two. This has been shown to be the case in electron micrographs (Penhoet (1967)) and also in the monoclinic form of aldolase (Eagles et al. (1969)). It was thought that the change in packing between the two forms caused by the change in precipitant, could best be examined by using models. Accordingly, model tetrahedra were made from table-tennis balls and a wire frame of the cell edges was constructed on a wooden base.

Unfortunately, the only information which was obtained from the models was that the molecules on the dyad in $P6_122$ could be similarly arranged around the hexads to those in $P6_222$. Thus, the packing could be the same. It was not possible, however, to derive any definite relationship between the overall packing in the two cells.

Phosphate ions are thought to bind in the active site of the enzyme (Kobashi et al. (1966), Morse and Horecker (1968)) which could cause a configurational change sufficient to alter the intermolecular binding (or associating) sites. This is unlikely to be the whole answer, however, some other effect being required as well. As has been mentioned already, aldolase crystals suspended in sulphate solution tended to be thixotropic whilst, in phosphate, this was not observed. This could be because the $SO_4^{=}$ is less efficient at neutralising some specific surface charge so that in the crystal, the molecules can attract each other more strongly causing a smaller unit cell.

It was hoped to be able to produce the monoclinic form from phosphate in order to compare the packing arrangement, as well as to collect data, but attempts to grow crystals at pH values between 7.0 and 7.6 either gave bipyramids or small needles except for the two tubes containing the sheaf-like plates the conditions for which could not be reproduced with the material available. This could be taken as evidence that the phosphate alters the surface charge in such a way as to render the plate form unstable within the range examined.

It is interesting, too, that the X-ray results reported here and those of Eagles et al. give rise to a molecular weight for the protein which is lower than the values obtained by either amino-acid analysis or the various hydrodynamic methods. This could arise from the incomplete blotting of the crystals in the protein content determination leaving some mother liquor adhering to the surface which would lead to a higher weight loss on drying than if all of the material was salt and water "diluted" with protein.

In conclusion, aldolase is a fascinating molecule not only from the point of view of its metabolic function but also from that of its own structure. The determination of the latter is already in hand with a search for suitable heavy atom derivatives. Its eventual solution will answer many of the questions which have been posed about its specificity, mechanism and active centre. There is one point of interest which concerns the resolution of the final map obtainable by X-ray techniques. Kochman, Penhoet and Rutter (1968) and Susor, Kochman & Rutter (1969) have been studying the heterogeneity of the molecule from muscle and find that the α_2-β_2 structure is not obligatory for aldolase A, but only an average value. Also, their hybridisation experiments seem to indicate that all types from α_4 to β_4 exist _in vivo_. If the distribution of subunits is random then they must be basically similar from an X-ray standpoint at low resolution, but the final high resolution structure determination will be clouded by this randomisation since all unit cells will not be identical thus causing un-evenness of packing between one cell and the next which, in turn will lead to an unusually high fall-off in intensity at large sine Θ values.

CHAPTER 5

Previous Work on β-Lactoglobulin

Introduction.

Milk is a substance of vital importance to mammals. It is composed of a large variety of compounds of differing sorts, each of which has interesting properties, both physical and chemical, in its own right and it is hardly surprising, therefore, that a great amount of effort has been put into the study of the various classes of compounds present in milk. Because bovine milk is so common, it is with this type that most of the work has been done. A general review of most aspects of one class of compounds in milk, the proteins, has been given by McKenzie (1967) and Table V. 1 gives his list of the proteins, excluding enzymes, found in mature bovine milk. This shows that the second most abundant protein is β-lactoglobulin. It is with this protein that the following work is concerned, and this review of the previous work is concerned mainly with those features which have a direct bearing on the structure and on the possible preparation of derivatives containing heavy atoms.

Separation from Milk

β-lactoglobulin was first isolated by Palmer (1934) and the protein accounts for some 10% of the total protein content in cows' milk. The separation and crystallisation of the protein is usually done by the method of Aschaffenburg and Drewry (1957a). Milk is warmed to 40°C and then treated with 200g./1. of milk of anhydrous Na_2SO_4 to precipitate the caseins, fats and globulins. The whey

Protein	Approximate Concentration (g./l.)	Percentage relative to total protein.
Caseins	25.0	80.4
β-Lactoglobulin	3.0	9.6
α-Lactalbumin	0.7	2.3
Serum Albumin	0.3	1.0
Immunoglobulins	0.6	1.9
Fat Globule Protein	0.2	0.6
Others	1.3	4.2

TABLE V.1 Bovine milk proteins, excluding enzymes, as given by McKenzie (1967).

proteins left in solution are further separated by the addition of HCl to lower the pH to 2.0 where α-lactalbumin and serum albumin are precipitated. β-lactoglobulin is left in solution and can be precipitated at pH 6.0 by the addition of a further 200 g./l. of $(NH_4)_2SO_4$. This allows concentration of the protein which can then be crystallised by dialysis against water. A modification of this method has been suggested by Armstrong, McKenzie and Sawyer (1967) which allows for a slightly fuller fractionation of the other species present and gives somewhat better yields.

Genetic Variants

In 1955, Aschaffenburg and Drewry showed that β-lactoglobulin can give two bands when electrophoresed at pH 8.6 and they separated and crystallised the two different types. Later, they went on to show that any cow has type A, type B or a mixture of types A and B (Aschaffenburg and Drewry (1957b)). From this, the synthesis of the protein was demonstrated by the good agreement obtained between calculated and observed frequencies of occurrence of the various types, as being controlled by two alleles of the one gene for which some cows are homozygous (either AA or BB) and others are heterozygous (AB). In 1962, Bell showed by means of starch gel electrophoresis that certain Jersey cows produced a form different to both A and B which he designated C. He found all six possible variations to be present in the herds which he examined and, where two types appear they do so in approximately equal quantities. Also, a daughter always has one type in common with her mother. A fuller account of the separation and crystallisation of this C variant has been given by Bell and McKenzie (1964, 1967).

Recently, a fourth genetic variant has been identified, and designated D, by Brignon et al. (1969) who quote a result found in 1966. This has been shown to have a glutamine residue in place of a glutamic acid residue in the sequence of the B chain and otherwise the properties appear to be identical with the other variant forms. A short length of sequencing around this change has been performed and the change occurs at residue 109 (see below).

Molecular Weight.

The molecular weight at the isoelectric point of the protein, which is around 5.2, has been found to lie within the range 35,000 - 37,000 by sedimentation - diffusion (Cecil and Ogston (1949)), light scattering (Halwer, Nutting and Brice (1951)), osmotic pressure (Bull and Currie (1946)) and X-ray crystallographic measurements (Green, North and Aschaffenburg (1956)). Other measurements made since these ones have been in agreement.

Subunit Structure.

An interesting point was raised by the work of Bull (1946) on the molecular weight as determined by the area of monomolecular films of the protein on a 23% A.S. solution. In the absence of metal ions the molecular weight was 18,000 but the addition of greater than $2x10^{-4}M$ Cu^{++} ions gave values of twice this. This was the first evidence that β-lactoglobulin was a dimer, and some time later Townend and Timasheff (1957) confirmed this by light scattering and ultracentrifugation in dilute solutions at pH values less than 3.5. Townend et al. (1961, 1964) and Timasheff and Townend (1961a) also confirmed this and, since then others have found this dissociation in acid solution. In the region between pH 3.5 and 5.4, β-lactoglobulin A has been shown to aggregate into a tetramer (i.e. eight subunits) by Timasheff and Townend (1961b), Witz, Timasheff and Luzzati (1964) and McKenzie, Sawyer and Smith (1967) using ultracentrifugation, electrophoresis, light scattering and X-ray small-angle scattering. The maximum tendency to aggregate is at pH 4.5, some 90% being tetramerised but variants B and C give no such results. The B and C forms barely associate although B has been shown to form some 30% of a mixed tetramer with A (Timasheff and Townend (1961b)). There is also some evidence of <u>dissociation</u> at low concentrations of all three forms between pH 3.5 - 9.2 (McKenzie and Sawyer (1967), Brignon et al. (1969)). Further, Townend et al. (1961, 1964) have shown that the dimers reassociate after dissociation only as homogeneous pairs, for example, as AA and BB but never as AB. Recently, however, Brignon et al. (1969) have obtained some evidence for the existence of hybrid BD forms.

<u>Titration of β-Lactoglobulins</u>

Linked to the dissociation/aggregation of the subunits are the changes in conformation as a function of pH (see below) and as a link between these are the studies on the titration of the protein. Tanford, Bunville and Nozaki (1959) and Nozaki, Bunville and Tanford (1959) found there to be an anomaly in that the amino-acids present in the analysis did not tally with those titrated unless the titration was performed after denaturation of the protein. They concluded that two carboxyl groups were being titrated abnormally and that this was associated with their being uncovered by the change in conformation between pH 6.5 and 8.0. Susi, Zell and Timasheff (1959) confirmed this by a study of the i/r spectrum. Using homogeneous protein, Tanford and Nozaki (1959) redetermined

	A	B	C	AB	AB (denatured)
α-carboxyl	2	2	2	-	-
Side-chain carboxyl (normal)	50	48	48	51	53
Side-chain carboxyl (abnormal)	2	2	2	-	-
Imidazole (normal)	4	4	4	6	4
Imidazole (abnormal)	-	-	2	-	-
α-amino	2	2	2	2	2

TABLE V.2 Groups titratable below pH 9.0 in β-lactoglobulin (36000). Data for A, B and C from Basch and Timasheff (1967) and that for AB from Tanford, Bunville and Nozaki (1959).

	A	B	C	D
Lys	15	15	15	15
His	2	2	3	2
Arg	3	3	3	3
Asp	16	15	15	15
Thr	8	8	8	8
Ser	7	7	7	7
Glu	25	25	24	25
Pro	8	8	8	8
Gly	3	4	4	4
Ala	14	15	15	15
Cys	5	5	3*	4*
Val	10	9	9	9
Met	4	4	4	4
Ile	10	10	10	10
Leu	22	22	22	21*
Tyr	4	4	4	4
Phe	4	4	4	4
Trp	2	2	2	2

TABLE V.3 Amino acid analyses of the genetic variants of β-lactoglobulin, A and B are from Frank and Braunitzer (1967), C is from Kalan, Greenberg and Walter (1965) and D is from Brignon et al, (1969), '*' indicates experimental inaccuracy.

their data for the A and B variants finding that A had two more non-terminal -COOH groups than B. This data was confirmed recently by Basch and Timasheff (1967) who also titrated the C variant. Brignon et al. (1969) showed by titration that the difference between the B and D variants is the replacement of glutamic acid by glutamine. Table V. 2 summarises the group counting for the four forms of β-lactoglobulin below pH 9.0.

Conformational Studies.

Tanford, Bunville and Nozaki (1959) showed that the specific rotation of β-lactoglobulin alters considerably and reversibly as the pH is increased from 6.5 to 9.0. This change is accompanied by the appearance of two carboxyl groups per molecule which are titrated as though they are imidazoles. The two forms, below and above the transition, were named N and R respectively by Tanford and Taggart (1961) who also showed from the form of the titration curve that these carboxyls are not in the area of contact of the two subunits. Dunnhill and Green (1965) showed that the free cysteine in each subunit is hindered in its reaction with pCMBS at low pH and Townend et al. (1969) and Pantaloni (1965) have showed that there is a hindered tyrosine residue also associated with the conformation change, although in this case it could be that the tyrosine is in the region of contact of the subunits and that dissociation in alkali is needed to free it for reaction. Pantaloni (1965) has suggested that the dissociation of the subunits, possibly triggered by the titration of the abnormal carboxyls, is responsible for the increased activity of the sulphydryl groups but Lyster (1964, per Dunnhill and Green (1965)) has

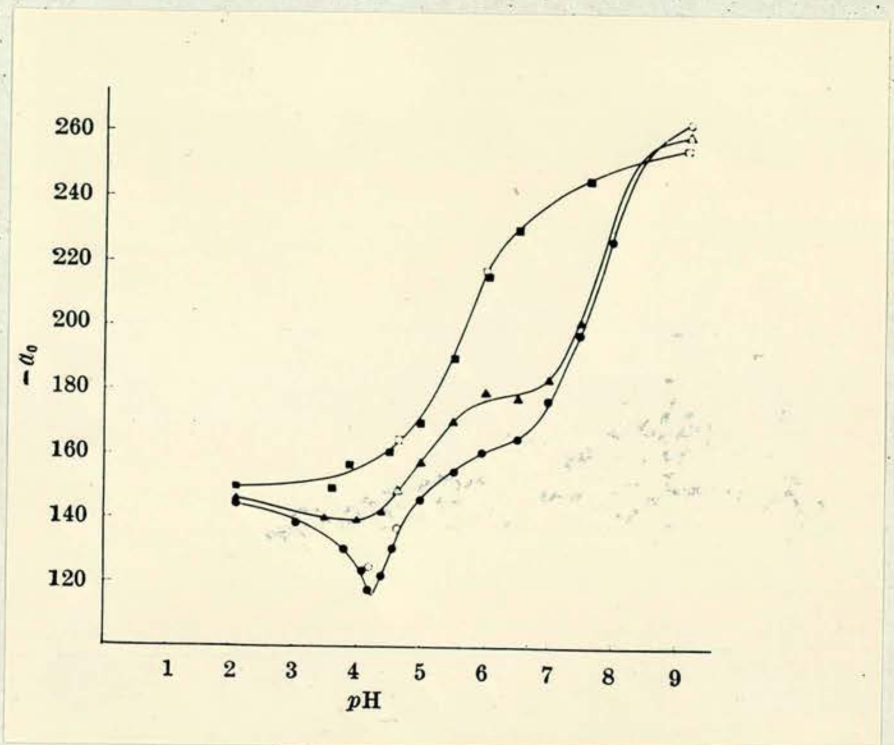


FIG. V.1 Data from McKenzie and Sawyer (1967) for a 10mg/ml protein solution showing the pH dependence of $-a_o$ at 20°C of the forms A (■), B (▲) and C (●) of β-lactoglobulin.

shown no increase in sulphydryl activity on dissociation below pH 3.5. At pH values above 8.5-9.0 the molecule dissociates and unfolds irreversibly and Pantaloni (1965) has described this in terms of the three forms:

$$N \rightleftharpoons R \longrightarrow S$$

Timasheff et al. (1966) and McKenzie and Sawyer (1967) have shown there to be two changes in conformation: the one above and another one between pH 4.5 - 6.0. This was shown by plotting $-a_o$, the term from the Moffitt-Yang equation (see p. 8) against pH (Fig. V. 1) and also titration data showed this lower change to be associated with the neutralisation of one proton in the case of the B and C variants and two in the A case. Thus, the conformational change in the lower pH range would appear to involve the region where the aspartic acid/glycine substitution occurs in the A and B variants. Townend (1965) and Basch and Timasheff (1967) also state that this aspartic acid residue in A could be the reason for the tetramerisation of A whilst no such aggregation is encountered with B, C or D. The R - S change has been found to occur irreversibly by alkali denaturation (Groves, Hipp and McMeekin (1951)), urea (Schellman (1958)), heavy metals (Pantaloni (1965)) and many others. It is associated with a large increase in the laevorotation usually followed by precipitation. There are conflicting reports, as has been said, as to whether the R - S stage involves merely unfolding of the polypeptide chain or this preceded by dissociation. The main states, however, can be written as

$$Q \underset{6.0}{\rightleftharpoons} N \underset{7.5}{\rightleftharpoons} R \underset{8.5}{\longrightarrow} S$$

In variant C, the ORD data show a difference occurring between pH 4.5 and 6.0 which Basch and Townend (1967) have put down to the uncovering of the histidine residue which is the difference between the variants A and B, and C. They also found that this group titrates anomalously (see Table V. 2).

Other studies on the conformation have been made with a view to elucidating the amounts of α-helix, pleated sheet and random coil present in the molecule. Before 1966, much work had been put into the interpretation of ORD data and a dilemma had arisen: the data seemed to point to a low percentage of helical content and yet a fairly highly ordered structure (Tanford et al. (1960)). With the advent of a means of "calibrating" the secondary structures, however, it became possible to interpret both ORD and CD curves with more accuracy (Sarker and Doty (1966) and Davidson et al. (1966)). Townend, Kumosinski and Timasheff (1968), therefore, were able to show by fitting observed and calculated ORD and CD curves for β-lactoglobulin that, for all three variants A,B and C, the relative percentages of α-helix, β-structure and random coil are something like 10%, 43% and 47% respectively. This was confirmed by the i/r studies of the protein in D_2O by Susi, Timasheff and Stevens (1968).

Amino-Acid Analyses and Amino-Acid Sequence.

The first amino-acid analyses of β-lactoglobulins A and B appeared independently in 1961 (Piez, et al., Gordon, Basch and Kalan). Since then several analyses have been performed on all variants. In Table V. 3, the data of Frank and Braunitzer (1967) are shown and are probably the most accurate since they are based

Leu-Ile-Val-Thr-Thr-Met-Glu-Lys-Gly-Asp-Leu-Gln-Ile-Lys-(Thr,Gly,Ala,Val,Trp)-

5 10 14

Tyr-Ser-(Ala,Leu)-Met-Leu-(Asx,Ser,Ser,Ala,Ala,Ile)-Leu-(Asp,Ala)-(Ser,Gln)-(Pro,

20 25 32

Ala,Leu)-Arg-Val-Leu-(Val,Leu)-(Asp,Asp,Tyr,Thr)-Lys-Lys-(Leu,Leu,Tyr)-(Glx,

40 50

Glx,Phe)-Cys-Met-Ser-(Asx,Ser,Glx,Glx,Glx,Pro-Ala,Leu)-ALA-Cys-Cys-(Val,Leu)-Arg-

57 68 70

(Asp,Thr,Pro,Val)-(Asx,Glx,Glx)-Ala-(Glu,Leu)-Lys-Asp-Phe-Lys-(Ala,Leu)-Lys-Ala-

81 85 90

(His,Pro,Met,Ile,Leu)-Arg-Lys-Leu-Leu-(Asx,Glx,Glx,Val,Ile)-Tyr-Val-Glu-GLU-Leu-

97 100 106 110

Gly-(Thr,Glx,Pro,Pro,Leu)-Lys-Asp-Glx-(GLY,Glx,Trp)-Cys-Ala-Gln-Lys-Lys-(Gln,Ala,

118 125

Ile,Ile)-Lys-Thr-Lys-(Pro,Ala,Val,Ile)-Phe-Lys-Leu-Asp-(Asx,Glx,Ala,Ile)-Asx-Lys-

132 140 147

(Glx,Leu)-(Asn,Ser,Gln,Pro,Leu,Phe)-Thr-Glx-Glx-Cys-His-Ile-COOH.

157 160 162

FIG. V.2 Partial sequence of the B-chain of β-Lactoglobulin after Frank and Braunitzer (1967). The peptide between 107 and 110 has the data of Brignon et al. (1969) included. Differences between variants, where known, are shown in capitals.

on the results of a partial sequence of the polypeptide chains.

A partial sequence of all of the amino-acids has been published by Frank and Braunitzer (1967). This is shown in Fig. V. 2 where the residues which are different between the A and B chains are marked. Brignon et al. (1969) have isolated and partially sequenced the peptide containing the difference between the B and D forms. This peptide, between residues 107 and 119 has been included in the sequence and the D variant has the glutamic acid residue 109 replaced by glutamine. Ambler and Scott (personal communication) have tentatively identified the free cysteine group as residue 123. The disulphide bridges are therefore either between residues 59 and 69, and 70 and 160 or between 57 and 70, and 69 and 160, the likelihood of a disulphide bridge between adjacent residues being somewhat remote.

The state of some of the amino-acid side chains in the protein has recently been discussed by Townend et al. (1969). Of the four tyrosine molecules per chain, two are readily available, one is somewhat hindered and the fourth is completely buried. The tryptophans are not involved in the association of the subunits and appear to be half buried although whether this means that one is available and the other buried or the both are half buried, is not clear from the fluorescence experiments. Alteration of the overall charge associated with the molecule by means of attaching groups to the cysteine via the sulphenyl iodide (Cunningham and Nuenke, (1959)) causes differences in the ability of the A variant to form tetramers suggesting that the cysteine group may be in the region of contact between the molecules when in the tetramer.

Variant and	Medium	pH	Type Code	Space Group	a (	b in Å	c)	α β	V (x10^{-5}) (Å^3)	Molecules (36K) Cell	A.U.
βA	H_2O	5.2	N	$P2_12_12_1$	68.6	70.2	137.8	-	6.63	8	2
βA+Zn,Cd,Pb	Ac	5.2									
βM	NaCl	3.5	P	$P6_3$	67.0	-	141.0	-	5.46	6	1
βA,βB*,βM	H_2O	5.2	Q	$P2_12_12_1$	69.3	70.7	157.5	-	7.74	8	2
βB	H_2O	5.2	R	$P2_1$	36.1	127.5	36.0	106° 05'	1.64	2	1
βB+Cd	Ac	5.2	S	$P2_1$	36.4	127.6	36.4	98° 12'	1.68	2	1
βB+Cd*	Ac	5.2	T	$P4_12_12$	69.2	-	138.8	-	6.65	8	1
βM	H_2O	5.2	U	$P4_22_12$	67.5	-	133.5	-	6.05	8	1
βM+MMA	A.S.	6.5	V	$P2_1$	35.1	56.2	72.4	102° 10'	1.40	2	1
βA+pCMS,IA	H_2O	5.2	W	$P2_1$	36.4	68.2	72.4	92° 12'	1.80	2	1
βA,βB,βM	A.S.	6.5	X	P1	38.1	49.7	56.6	122° 45', 97° 31', 104° 05'	0.83	1	1
βA,βB,βM	A.S.	7.6	Y	$B22_12$	55.7	67.2	81.7	-	3.06	4	1/2
βA,βB,βM	A.S.	7.6	Z	$P3_221$	54.4	-	113.1	-	2.89	3	1/2
Buffalo	H_2O	5.2	ØR	$P2_1$	35.9	127.7	35.9	106° 17'	1.63	2	1

TABLE V.4 Crystal forms of β-Lactoglobulin based on data from Aschaffenburg et al. (1965).

Abbreviations: βA, βB and βM are the pure A, pure B and mixed genetic variants respectively.

βB* is a more stable crystal form of βB.

H2O refers to crystals prepared by dialysis against water.

Ac refers to those crystals obtained from dilute acetate without dialysis.

pCMS and IA refer to p-chloromercuribenzene sulphonic acid and iodoacetamide, respectively.

Note: The hand of the 4_1 axis has not been established.

Cell dimensions are accurate to 0.5 %.

Finally, because there is almost no change in the specific rotation when the two C-terminal residues are removed with carboxypeptidase, it is suggested that this terminus must be on the surface.

Crystallographic Studies of β-Lactoglobulin

Although crystals of β-lactoglobulin were obtained by Palmer (1934), it was not until 1956 that the A and B variants were examined by X-ray crystallography along with the mixed form (Green, North and Aschaffenburg (1956); see references therein to the earlier X-ray work). These crystals were prepared by dialysis. A fuller description of the crystal forms of the protein, both separated and mixed, is given by Aschaffenburg, Green and Simmons (1965) including the forms obtained by salting out. Table V. 4 gives the data of Green and his coworkers brought up to date by the inclusion of Lattice V parameters obtained by Komorowski (1971).

As was mentioned in Chapter III, salt-free crystals are not always of much use for the production of heavy atom derivatives since the inclusion of heavy atom salt ions tends to salt in the protein into solution again. If this does not happen then distortion of the lattice often occurs. Consequently, the use of salt-free crystals for the determination of the structure has not been very successful (Aschaffenburg et al., (1965)).

Five forms of crystals have so far been grown at near neutral pH by salting-out. Each of lattices X, Y and Z has been the subject of preliminary and low resolution X-ray investigations, electron density maps to 6Å having been calculated, and it is with these three forms with which the remainder of this chapter will be concerned.

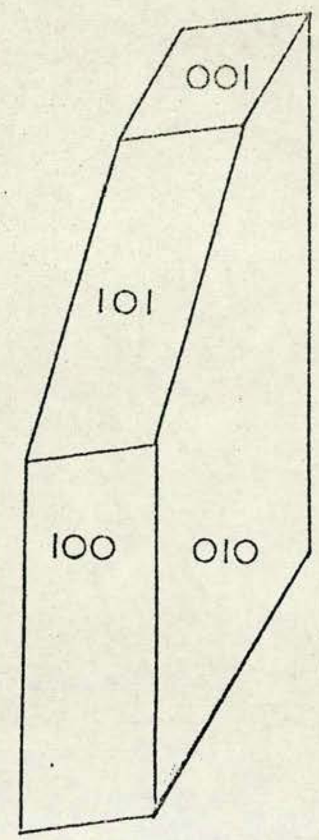


FIG. V.3 Typical lattice X crystal.

Lattice X.

Much of the work on this form has been done recently by Komorowski (1971). The crystal habit is sketched in Fig. V. 3, and the space group is triclinic, P1, there being one molecule in the unit cell. This means that there is a non-crystallographic twofold axis relating subunits. Crystals are stable up to pH 6.9 but are generally prepared between pH 6.0 - 6.5 from 2.0 M ammonium sulphate with a final protein concentration of some 25mg./ml. in the crystallisation tube.

It is thought that the X form corresponds to the lower conformation as obtained by Tanford, Bunville and Nozaki (1959) and designated N. Dunnhill and Green (1965) showed that in this conformation the protein's free sulphydryl is less reactive towards mercurials such as pCMBS. Nevertheless, pCMBS gives an isomorphous derivative useful to at least 6Å resolution. Komorowski found however, that, whilst he could obtain the heavy atom parameters for the $HgI_4^=$ derivative from the projected $(\Delta F)^2$ syntheses, those of other derivatives were uninterpretable. This was partly because there was no centrosymmetric projection so that Fourier maps could not be obtained from the phases calculated from the one, interpreted vector map to aid the interpretation of the separate $(\Delta F)^2$ syntheses. The derivatives tried were those known to attach to the other forms: e.g. MMA, $Pt(NO_2)_4^=$, $Au(CN)_4^-$ etc.

This problem was finally overcome by collecting full three-dimensional sets of data for the derivatives $Pt(NO_2)_4^=$, pCMBS and $HgI_4^=$. $(\Delta F)^2$ maps were found to be interpretable and the parameters

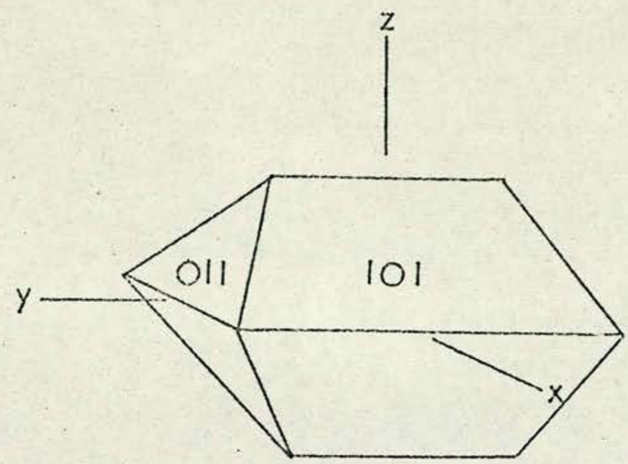
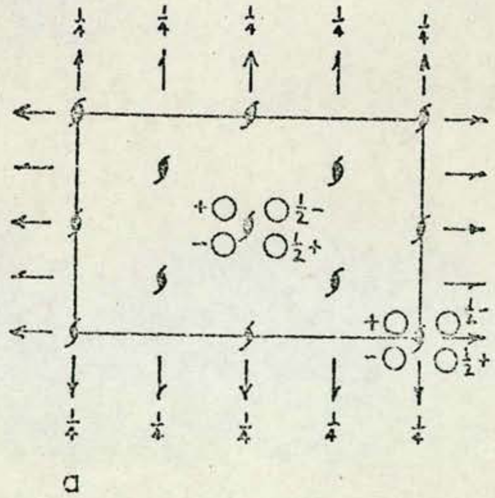


FIG. V.4 Typical lattice Y crystal.



Origin at 12_12

FIG. V.5 Space group $B22_12$.

obtained, once refined, enabled a protein Fourier map to be calculated. The map, though uninterpretable as a whole had certain features in common with those of the other forms, especially near to the sulphydryl sites.

Lattice Y.

Lattice Y appears to be the more stable of the two high pH forms of β-lactoglobulin. The typical habit is sketched in Fig. V. 4 and in Fig. V. 5 the space group, $B22_12$, is shown with the eight general positions. The unit cell contains four molecules of protein which implies that the molecular dyad is also a crystallographic one. Since these crystals can be obtained with the mixed protein as well, and since the difference between the A and B chains is limited to two amino-acids, only as high resolution is approached will "blurring" of the protein Fourier map in the regions of these changes become appreciable.

Crystals are stable down to pH 7.1, being prepared at about pH 7.6 from approximately 2 M A.S. and a final protein concentration of 10-30 mg./ml. Occasionally lattice Z crystals appear in these preparations and, in some cases, in the same tube. It is possible that the Z crystals, which can be found to convert to Y on standing, are the more stable kinetically and hence can be formed first, the thermodynamically more stable Y form taking a longer time to grow. The energy barrier between these two forms must, in any case, be small.

Green and his coworkers (unpublished) have made extensive diffusion trials on lattice Y. In general, it was found that at 10mM concentration of heavy atom (a heavy atom to protein subunit ratio of about 2.5 : 1) either nothing significant occured within about six weeks or else very large changes accompanied by distortion

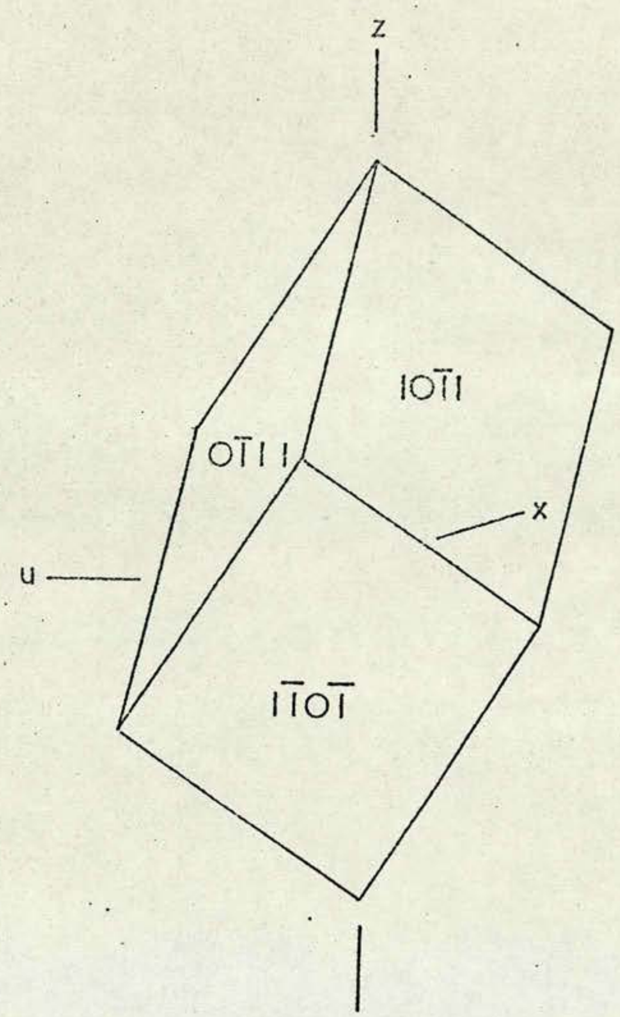


FIG. V.6 Typical lattice Z crystal.

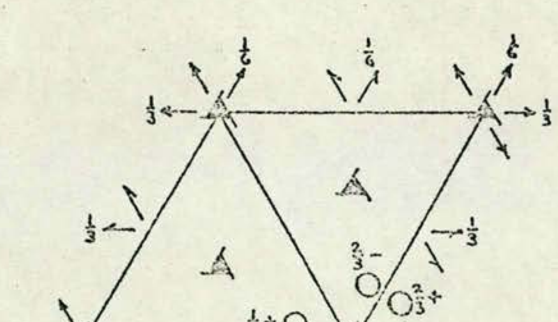


FIG. V.7 Space group $P3_{2}21$.

of the lattice were obtained. However, three derivatives have been prepared and used to calculate a protein electron density map. These were MMA, $Pt(NO_2)_4^=$ and $HgI_4^=$ and they provide a starting set of α_P values for screening further derivatives. One unfortunate feature of the Y lattice is that the z coordinates of the heavy atom sites are rather close to $z=\frac{1}{4}$. This introduces a spurious centricity into the phasing because small displacements from $z=\frac{1}{4}$ are ambiguous. Another rather awkward feature of the Y crystals is that $HgI_4^=$ is one of the heavy atoms used. If this complex is to be used for high resolution studies it is essential that all of the I^- ions are positioned, as well as the mercury atoms.

Lattice Z.

Lattice Z crystals are the less stable form obtained from approximately 2M salt solution as described above. Fig. V. 6 shows the habit of typical crystals and the space group is $P3_221$, shown with the six general positions in Fig. V. 7. There are three molecules in the unit cell which, once again, implies a coincidence of the molecular and crystallographic dyads. The hand of the screw has been determined by anomalous scattering measurements to be 3_2 rather than 3_1 as was originally supposed (Simmons (1965)).

This form looks slightly more promising for high resolution study, mainly because there is no ambiguity about the values of the coordinates. It has a centrosymmetric projection, [010] from which all three coordinates can be obtained. The derivatives used were MMA, $Pt(NO_2)_4^=$ and $HgI_4^=$ the latter one being rather poor

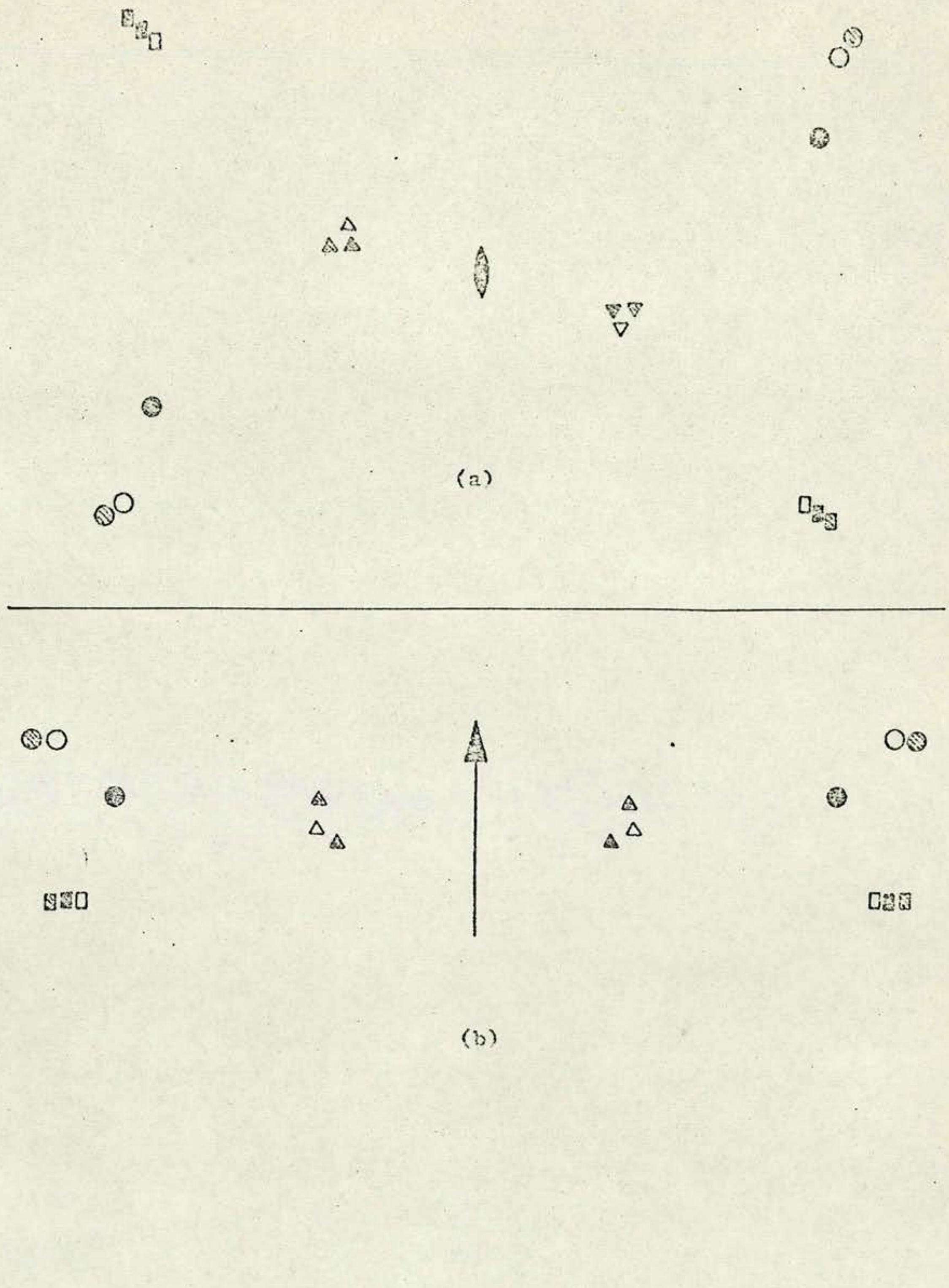


FIG. V.8 Projections (a) along and (b) perpendicular to the molecular dyad showing the heavy atom arrangements in lattices X,Y and Z. (◍): the sulphydryl site, (▲): the NGI site and (▦): the PTN site. The solid symbols are for lattice X, the hatched for lattice Y and the open for lattice Z. The scale is 1cm = 2Å.

	LATTICE Y			LATTICE Z		
	x	y	z	x	y	z
MMA	0.239	-0.039	0.240	0.854	0.371	0.093
KGI	0.573	0.037	0.263	0.102	0.448	-0.002
	-	-	-	0.984	0.162	0.156
PTN	0.367	0.337	0.301	0.259	0.485	0.019

TABLE V.5 Fractional coordinates of the equivalent heavy atom sites in lattices Y and Z.

because of high fall-off and differing amounts of substitution within crystals from the same preparation. Despite this drawback, a protein Fourier map has recently been produced which, although uninterpretable as a whole does look promising, with rods of density which might be α-helix. However, higher resolution maps will have to be produced before the peptide backbone can be followed with any degree of certainty. This is not in anyway unusual since, unless there is a large amount of helical content as in myoglobin for instance, interpretation of the 6Å map is unlikely to be possible. Blow and Steitz (1970) give a table of the stages of interpretation of the various proteins whose structures have been determined.

Comparison of the Three Forms.

At 6Å resolution, by far the most interesting point which emerges is that concerning the positions of the heavy atom sites relative to the molecular dyad (Green et al., (1971)). Since the three successful derivatives in each form have been the same, except that the "thiophilic" reagent in X was pCMBS rather than MMA, this provides a possible means of comparing the structures of the three forms. Fig. V. 8(a) and (b) shows the projections along and perpendicular to the dyad, determined by assuming that the structure is invariant with pH for the HGI and PTN* sites. Going from X to Y and Z, it can be seen that the MMA site moves relative to the others and this is consistent with the change of conformation with an increase of pH. Table V. 5 lists the fractional coordinates of the various sites in lattices Y and Z. It will be noticed that there is an HGI minor site in Z which is not occupied in Y suggesting

---

* The sites referred to are those at which $HgI_4^=$ (HGI) and $Pt(NO_2)_4^=$ (PTN) and MMA (MMA) are found to exist in the various crystal forms.

that the difference in packing may be affecting the availability of this minor site.

One final point seems clear. The diffusion trials in lattices Y and Z by Green and his colleagues and lattice X by Komorowski (1971) have shown β-lactoglobulin to be capable of forming derivatives. These have tended to be accompanied by a large amount of fall-off and a certain amount of distortion of the lattice. Other trials have failed in that nothing seems to have happened to change the diffraction spectrum. What is required is an intermediate stage where moderate changes occur but with very little distortion rather than the all-or-nothing state which seems to exist.

The full tertiary structure will clarify most of these points and, since forms at various pH values are available for X-ray work some idea of the changing conformation in solution should be gained.

CHAPTER VI

β-Lactoglobulin and Tetracyanoaurate (III)

SECTION I

Introduction

As has been mentioned in the previous chapter, the heavy atom derivatives used for the high pH forms were not all satisfactory for high resolution studies. The MMA derivative is a good one but the HGI one suffers from two major drawbacks. First, for higher resolution work than 6Å, the positions of the iodine atoms would have to be found since they have a considerable scattering power. What makes matters worse is that it is not certain what form the complex takes in its binding to the protein: whether it is HgI_2, HgI_3^- or $HgI_4^=$ and this could conceivably vary from preparation to preparation and even from crystal to crystal. Second, there is an increasing degree of non-isomorphism which becomes apparent at resolutions greater than 6Å. Thus, on two counts, $HgI_4^=$ is not a very suitable 'heavy atom' for high resolution studies.

The following work was undertaken with this in mind and it was hoped to be able to produce some derivatives which could be used for a high resolution study of the structure. At the same time it was hoped that it might be possible to find a more systematic way for preparing such derivatives which could be applied not only to β-lactoglobulin but also to other proteins. It was considered that solution studies might help considerably in this latter aim by showing if binding of a specific nature was occurring in solution. If such specific interactions were taking place and could be identified, then could it be shown that the binding site was the same in the crystal? This in turn might aid the interpretation of the protein electron density map by allowing the fitting of the amino-acid sequence, whether partial or complete, if it were known to which residues the heavy atom was binding.

Studies of proteins can conveniently be made in solution by visible/ultraviolet spectroscopy and polarimetry. This latter means might be expected to be of more than normal interest in the case of β-lactoglobulin because of the marked conformational change occurring as the pH is raised from 6 to 8 (Tanford, Bunville and Nozaki (1959)), hereafter referred to as the 'Tanford transition'. It was with these two methods of monitoring any possible reaction which might occur in solution that the following studies were undertaken.

Studies with Gold Compounds.

$HgI_4^{=}$ has been mentioned as being unsuitable for further studies but what of the lighter tetra-coordinate compounds of mercury? The chloride does not exist to any appreciable extent in solution as $HgCl_4^{=}$ and $HgCl_2$ has been found to go to the MMA site, that of the free sulphydryl group (Green, unpublished work). Likewise, $Hg(CN)_2$ appears to go to this site and it also causes the c axis to shorten in Lattice Y (Green, unpublished work). There was also some doubt as to the exact nature of the compound which was binding and causing the intensity changes. It appeared, then, that mercury compounds of this type could be discarded.

The next lighter element in the periodic table was gold and it was decided to examine the various light-atom complexes of Au(III). Preliminary diffusion trials carried out by Dr. Green and his colleagues at the Royal Institution, London had shown that the

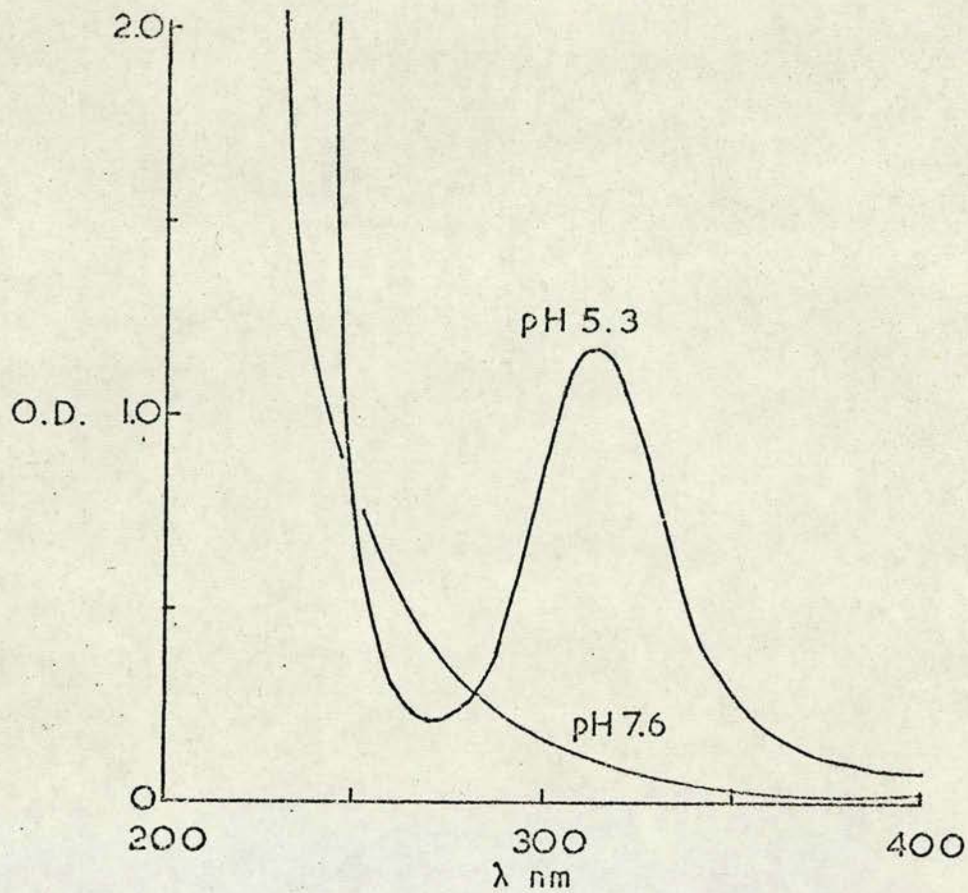


FIG. VI.1 Spectra of tetrachloroaurate(III) in acid and basic solution. The complex was 25μM.

tetracyanoaurate (III) ion (TCA) did in fact cause changes in lattice Y crystals. These were thought to be encouraging enough to warrant further investigation. Consequently, crystals had been prepared with stoichiometric amounts of TCA and brought to Edinburgh.

Inapplicability of Tetrachloroaurate (III).

Sodium tetrachloroaurate is readily obtainable commercially in a pure form. It is easily hydrolysed in basic solution according to the stepwise equations:

$$AuCl_4^- \underset{Cl^-}{\overset{OH^-}{\rightleftharpoons}} AuCl_3OH^- \underset{Cl^-}{\overset{OH^-}{\rightleftharpoons}} AuCl_2(OH)_2^- \underset{Cl^-}{\overset{OH^-}{\rightleftharpoons}} AuCl(OH)_3^- \underset{Cl^-}{\overset{OH^-}{\rightleftharpoons}} Au(OH)_4^-$$

These equilibria have been examined by Bardet and Bontoux (1961) who have shown that at about pH 7.5 there is almost no $AuCl_4^-$, about equal quantities of $AuCl_3OH^-$ and $AuCl_2(OH)_2^-$ totalling about 95% of the species present, the remainder being mostly $AuCl(OH)_3^-$. Therefore, in the solutions in which lattice Y crystals grow, the tendency is towards the righthand side of the equilibria. This was observed almost immediately the complex was dissolved in 0.05M phosphate buffer or 0.5M NaCl solution at pH 7.6. The yellow colour disappeared over a period of hours and the u/v spectrum changed considerably. The spectra at pH 5.3 and pH 7.6 are shown in Fig. VI. 1.

Kendrew (1962) has discussed the substitution of the chlorides of $AuCl_4^-$ in the same way as the hydroxide above, by NH_3 from the solution of ammonium sulphate used to grow myoglobin crystals. This was to explain the association of $AuCl_4^-$ with a histidine residue in myoglobin crystals. The same situation might arise in β-lactoglobulin so that a series of differently charged ions which could bind to various sites on the protein would be formed. This also would be apt to cause a 'messy' derivative with several partially occupied sites, the occupancies of which would be likely to alter from one preparation to the next.

Further, it was found that a thiol could readily reduce the $AuCl_4^-$ to Au, an observation borne out by the oxidation potentials for gold chlorides, that of the oxidation of a thiol being about +0.2V (see below).

$$\begin{tikzcd} AuCl_4^- \arrow[r, "+0.926V"] \arrow[rr, bend right, "+0.991V"'] & AuCl_2^- \arrow[r, "+1.14V"] & Au \end{tikzcd}$$

pH 7.0, phosphate buffer (Bardet & Bontoux (1961))

In crystalline β-lactoglobulin, reduction of the chloroaurate to colloidal gold has been observed, the crystals turning a reddish purple (Green and Komorowski, unpublished work).

Tetrachloroaurate, therefore, was discarded on account of its easy reduction and solvolysis. Some means of preventing these phenomena and yet not adding bulky or heavy ligands was required. Replacing the chloride by cyanide stabilises the complex, had already been found to cause large changes to the native diffraction pattern and was therefore the obvious choice.

The Tetracyanoaurate (III) Ion.

Britton and Dodd (1935) have examined the titration of $NaAuCl_4$ solution with KCN both conductimetrically and with a glass electrode. The results of this latter titration appear in Fig. VI. 2. As can be seen, at pH 8.0 there are four cyanides per gold ion and analysis of the solution at this pH showed there to be no free cyanide and only 0.26% hydrolysis. They also found the combined stability

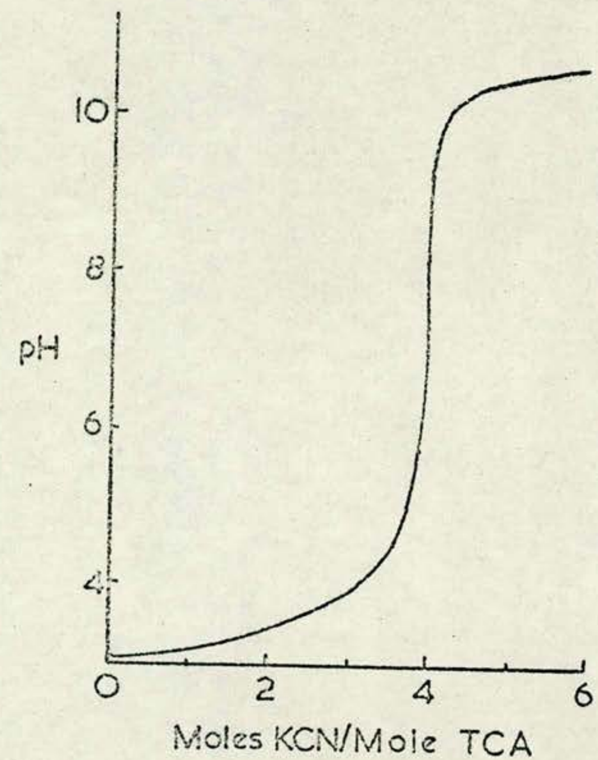


FIG. VI.2 Titration of tetrachloroaurate(III) with KCN (after Britton and Dodd (1935)).

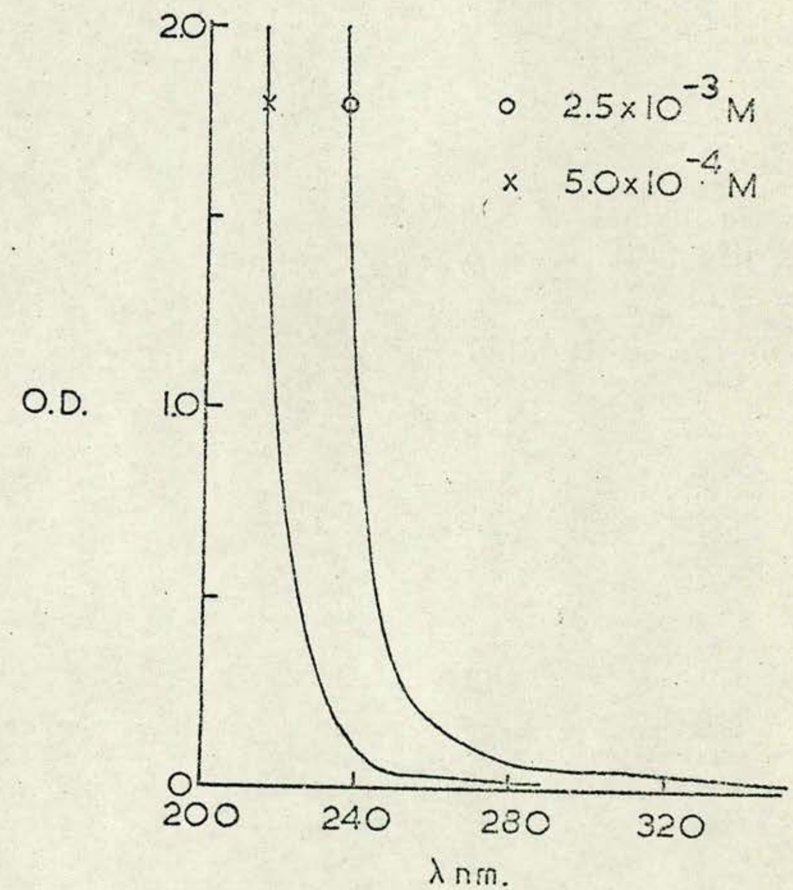


FIG. VI.3 Spectrum of a 25mM solution of TCA.

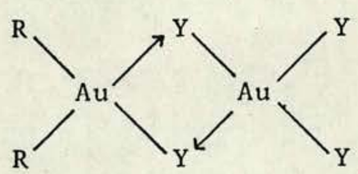
constant, β_4, to be about 56 (i.e. $\log K_1 + \log K_2 + \log K_3 + \log K_4 = 56$). The u/v spectrum is shown in Fig. VI. 3.

Gold (III) has a d^8 configuration and therefore forms square planar compounds unlike the Hg (II), d^{10} ones which are tetrahedral. Bertinotti and Bertinotti (1968, 1970) have determined the structure of the TCA group by both X-ray and neutron diffraction. The group is nearly planar with the following dimensions:

Au-C : 1.98 ± 0.01Å ; C-N : 1.14 ± 0.02Å ; Au-C-N : $177 \pm 2^{\circ}$

The nitrogens are also hydrogen-bonded to water molecules in the crystal, the N-O' separation being about 3Å. The cyanides' ability to form hydrogen bonds is a useful feature as far as binding to proteins is concerned.

However, the different configuration of TCA from that of the $HgI_4^{=}$ tetrahedral ion meant that binding might occur at a different site. Also, bridged compounds of the type shown below are possible (see, for example, Durrant and Durrant (1962)).



R = Alkyl
Y = Cl^-, Br^- or CN^-

Might this type of complex form with the protein? The first stage was to find the binding site by X-ray methods.

Two Dimensional X-Ray Work

Because lattice Y is orthorhombic and belongs to point group 222, it has three, mutually perpendicular, centrosymmetric projections. It was hoped, therefore to be able to determine the site or sites of substitution from projections down each of the three major axes. Crystals of lattice Y had been prepared by Dr. Green and his coworkers in London and treated in the following manner. The mother liquor from the native preparation was removed, treated with 1.3mls. 4M A.S. to take the final concentration of sulphate to 2.5M and made 50mM in KCN. This had no visible effect on the crystals when they were reimmersed in the treated mother liquor. $NaAuCl_4.2H_2O$ was then dissolved in the mother liquor once more removed from the tube and then added back slowly over a period of days. This gave a final molar ratio of TCA to protein of 4:1, earlier trials having shown that ratios of 2:1 and 3:1 had only caused small intensity changes. After some weeks, some of the crystals had cracked indicating that the TCA was having some effect.

10° precession photographs were taken using Ni filtered Cu Kα radiation, along each of the three principal axes. These showed large changes not only at low values of sin θ but also further out. These were coupled with a larger amount of fall-off than was found in the native crystals and, what was more unfortunate, changes in the cell dimensions. The unit cell parameters are shown below together with those of the native.

$\alpha = \beta = \gamma = 90^{\circ}$

	Derivative	Native
a:	55.7)	55.7)
b:	67.9) $\pm$0.5Å	67.2) $\pm$ 0.5Å
c:	78.4)	81.7)

The most marked change was that of a 4% shortening in the c axis, which is also the molecular dyad. Was there some kind of intermolecular oligomer being formed causing the molecules to pull themselves together? This external lack of isomorphism meant that the data could not possibly be used for phase determination. However, it was decided to attempt to determine the heavy atom coordinates for two reasons. First, to find out where the TCA was binding in relation to the other known sites in an attempt to explain why the distortion should occur principally in the one direction. Second, appreciable anomalous dispersion effects were observed in a [10$\bar{1}$] photograph. It was hoped that the lack of isomorphism would not be so large in lattice Z so that the derivative, if found to bind in a suitable (new) position, could be used, along with the anomalous scaterring, to extend the resolution beyond 6Å in the other high pH form.

Before proceeding, however, it was considered worthwhile to check what effect the lack of isomorphism had on the intensity distribution compared with that of the native. Crick and Magdoff (1956) suggested that an examination of the function

$$\Phi \Delta I = ((\overline{\Delta I})^2)^{\frac{1}{2}} / \bar{I}_p$$

over a series of small ranges of sin Θ, or, for centric reflections,

$$\Phi' \Delta I = ((\overline{\Delta F})^2 / \bar{I}_p)^{\frac{1}{2}} \doteq \tfrac{1}{2}\Phi . \Delta I$$

Carlisle and Palmer (1962) have applied this function to ribonuclease and its derivative with p-chloromercuribenzoate to show the unsuitability of the derivative to a resolution greater than 6Å. If the only changes in the derivative arise from the inclusion of a heavy atom, the function $\Phi\Delta I$ should have a similar shape to the scattering factor curve for the heavy atom. If changes occur in the position and shape of the protein as well, then there will be an increase in $\Phi\Delta I$ at higher values of sinΘ. A complete, three-dimensional set of native lattice Y data had been collected by Green et al. (unpublished, 1963) to a resolution of 3.3Å. It had also been

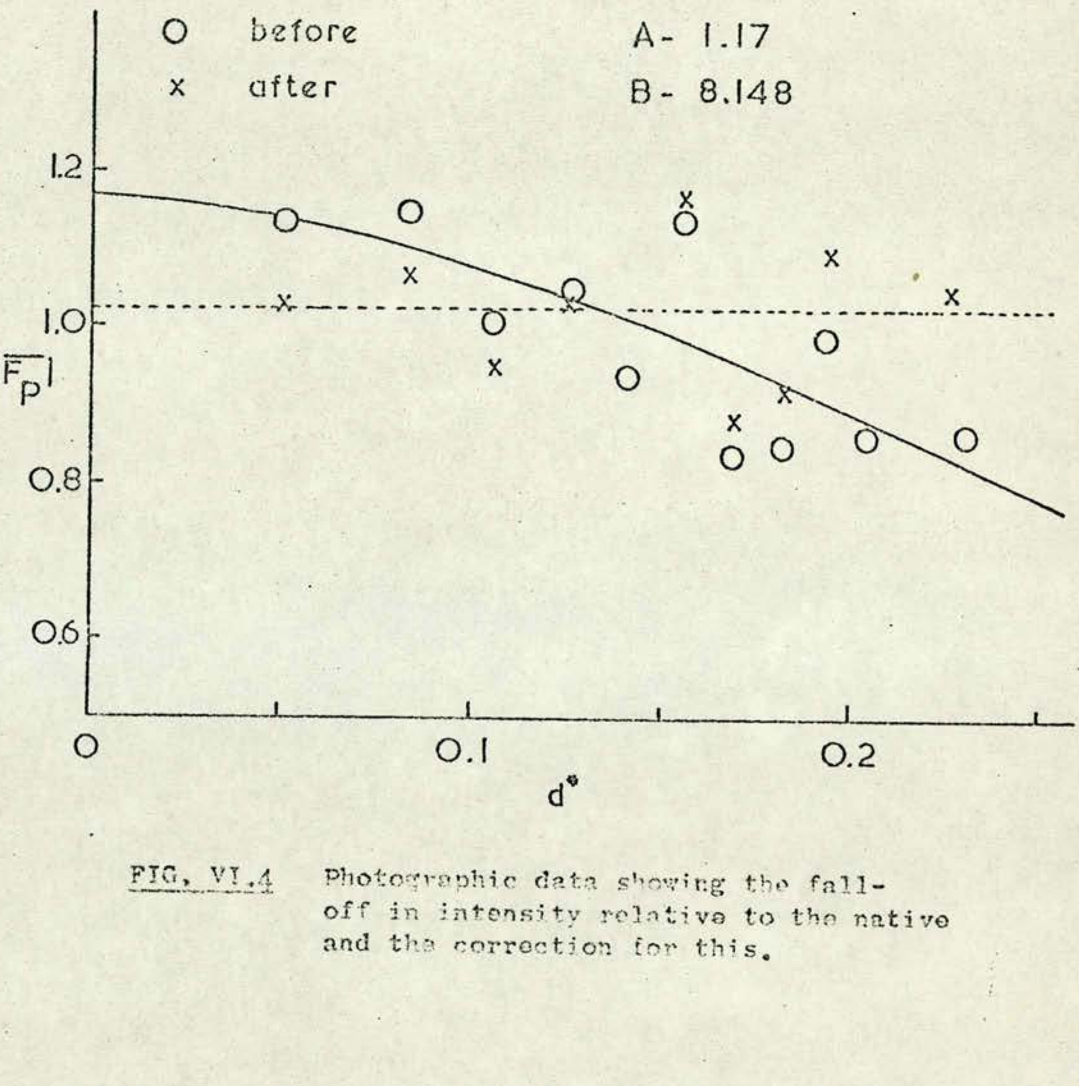


FIG. VI.4 Photographic data showing the fall-off in intensity relative to the native and the correction for this.

set on an approximately absolute scale. This set was available for scaling and comparison purposes.

Processing of the Photographic Data.

The photographs were measured on a Joyce-Loebl Mk IIIB twin beam microdensitometer using a light beam about 1/3 by 1/10 of the diameter of the spot, the smaller dimension being in the direction of the scan. The raw intensities were obtained then by measuring the peak height. The values thus obtained were corrected for Lorenz and polarization effects using a programme written by Mr. E.S.Komorowski for the KDF 9 computer. The data from each photograph in turn were then scaled to the native set using a programme called COLLATOR.

$$F_{scaled} = k.F_{measured} \quad \text{where } k = \sum |F_P| \,/\, \sum |F_{PH}|$$

A comparison of the values of $\sum' |F_P| \,/\, \sum' |F_{PH}|$ where the summations are done over all of the data contained in annuli of approximately equal $\sin\theta$, then showed that a scale factor of the form

$$k \;=\; A.\exp(-B.4\sin^2\theta/\lambda^2)$$

would be more appropriate. The values of A and B were determined for all photographs together since any one photograph had only three or four values on which to base the calculation and because an examination of these values showed there to be similar fall-off on all of the photographs. The values of A and B were determined by a least-squares procedure written for the purpose and run on the departmental PDP 8 computer. The scale factor was applied by use of a short program, FALLOFF, to the lists produced by COLLATOR. These data, together with the values of A and B, showing the effect of the scale factor, are shown in Fig. VI. 4. This correction allows for the effect which the heavy-atom complex has on the protein crystals on its percolation, binding and consequent disarrangement of the lattice. This disarrangement should be isotropic and, as far as could be seen, it was. In effect, it takes into consideration the difference in temperature factors between native and derivative. Thus, an isotropic scale factor taking into account this effect was considered justifiable.

Having scaled the data sets in the above manner, it was then necessary to scale them together to make best use of the repeated measurements on the h00, 0k0 and 00k rows. The method of scaling used by COLLATOR required there to be at least one reflection in common and a basic set to which the others are scaled. The method has been compared with the conventional least-squares method of Hamilton, Rollett and Sparks (1965) and found to give scale factors differing by about 1% (Simmons (1965)), provided there are sufficient terms in common.

In space group $B22_12$, neglecting anomalous scattering, hkl = $\overline{hk}l = \bar{h}k\bar{l} = h\overline{kl} = \overline{hkl}$ or only a quarter of a zero level photograph has independent reflections. This meant that each intensity was measured four times, the average being taken as the true intensity thus cutting down the amount of data it was necessary to store on magnetic tape. However, the space group's having B face-centring causes all reflections with h+l = 2n+1 to be absent and this in turn causes there to be fewer common reflections for scaling purposes. To overcome this problem, and to minimise any effect the order of scaling the photographs might have on the individual scale factors whilst building up a set with all terms included, a procedure

Photograph	Scale Factor		
	1st Cycle	2nd Cycle	3rd Cycle
[010]	-	2.265	2.279
[100]	0.819	0.833	0.036
[$10\bar{1}$]	2.529	2.547	2.532
[001]	2.174	2.151	2.165
Mean D %	12.9	11.8	11.9
Mean scale factor change	-	1.2	0.45

TABLE VI.1 Scaling together of the photographic data sets showing the convergence.

incorporated in COLLATOR was used. A store resembling the three-dimensional lattice was set up in the computer and each set, having been scaled to an arbitrarily chosen basic one, was added to the set (or sets) already in the store, averaging the common reflections and keeping a record of the number of times a reflection has been measured. Having added all of the sets of data, each set in turn was removed, recompared, rescaled and added in once more. The process was then repeated. It was found that this altered the separate scale factors after the first cycle by less than 0.5%. The results are shown in Table VI. 1. After the final cycle, the store was output to magnetic tape as an indexed list of h,k,l, $|F_{PH}|$ and the number of times the reflection has been measured.

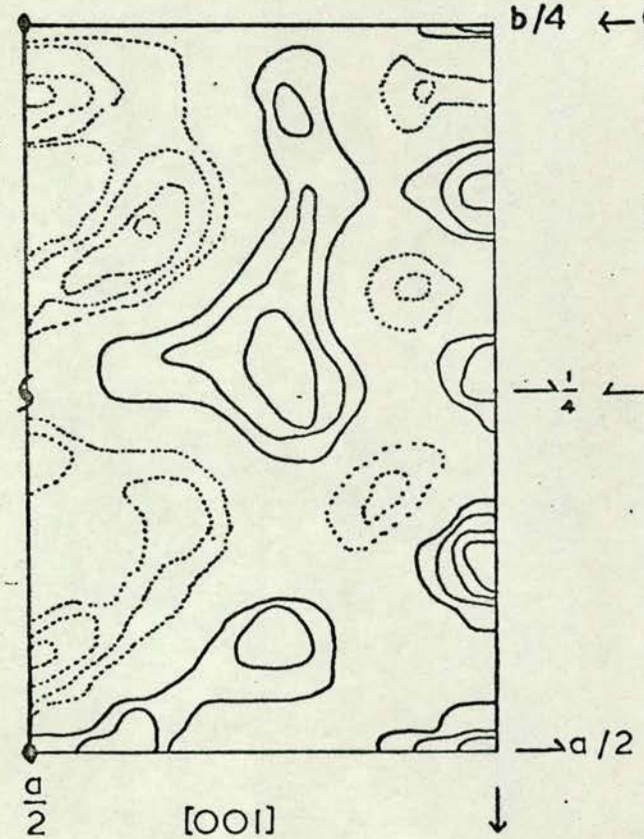
The differences between the various measurements of the same reflections can be combined to give some idea of accuracy of the data. COLLATOR prints out a radial distribution of values of

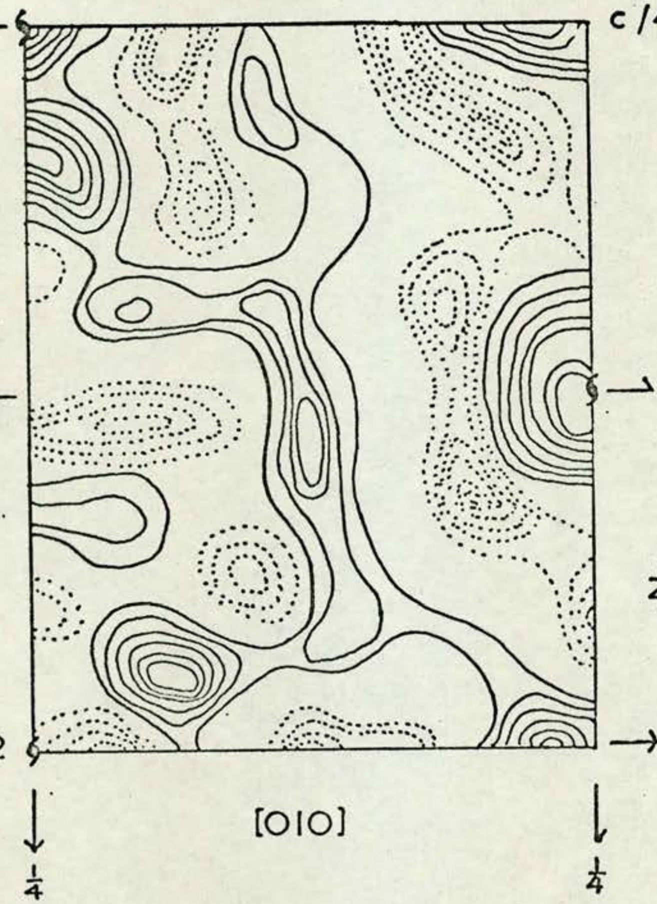
$$\sum_N |F_1| / N \ , \ \sum_N |F_2| / N \ \text{and} \ \sum_N |F_1 - F_2| / N$$

where N is the number of common reflections in the annulus between those stored (F_1) and those being compared (F_2). Thus the reliability of a complete set of data can be defined as

$$D = \sum_N (|F_1 - F_2|). 100 / \sum_N |F_1|$$ (Drenth et al. (1965), Scouloudi (1969))

Typical values of D for proteins are between 5-8% (see Simmons (1965), Drenth et al. (1965), Blow, Rossman and Jeffrey (1964) and Scouloudi (1969)) for both photographic and linear diffractometric data. The mean value of D for all photographs for the TCA-β-lactoglobulin derivative was 11.8% which was rather high but similar to that for the HGI derivative in lattice Z and also to that of





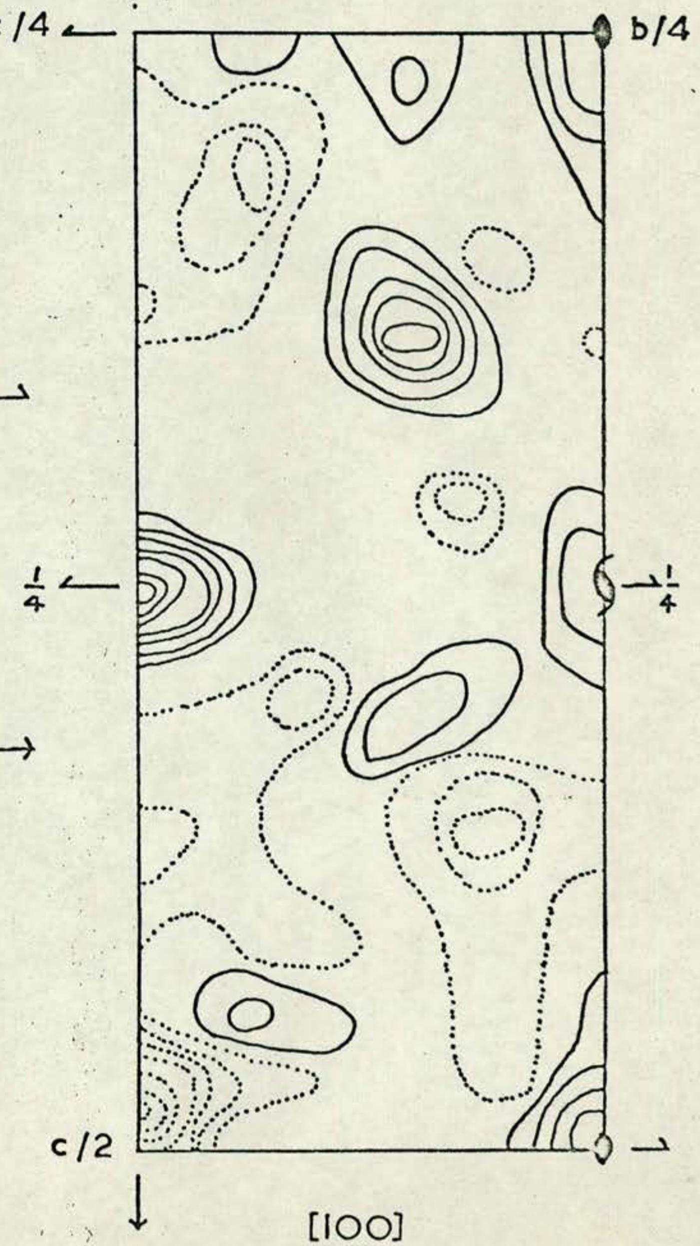


FIG. VI.5 The three difference Fourier projections, [100], [010] and [001] of the TCA derivative in lattice Y, with contours drawn at 1 electron/Å^2 intervals.

	x	y	z	Peak Heights [100]	[010]	[001]
Positive Peaks						
	0.25	0.25	0.25	189	634	184
	0	0.25	0	231	239	221
	0.24	0.15	0.15	353	234	183
	0.49	0.05	0.24	392	704	155
Negative Peaks						
	0.48	0	0.47	-361	-197	-225
	0.36	0.17	0.10	-168	-195	-172
	0.18	0.03	0	-175	-219	-207

<u>TABLE VI.2</u> Coordinates of peaks, both positive and negative, which are consistent with the three projections of the TCA derivative.

simply by altering the phases of those reflections having k odd to $\pm\ \pi/2$ instead of 0 or π. It was clear that any information obtained from these maps was likely to be dubious. Interestingly, however, the $HgI_4^=$ site did appear, although in one projection, the [001], it was the smallest peak.

Why were the maps so "messy"? There were several possible reasons, the real answer probably being a combination of these. First of all, the set of phases (or signs) used were calculated from the three derivatives, MMA, PTN and HGI and it was thought possible that these sites might appear as 'ghosts' in a map calculated using these signs. In other words, the phases might be biased towards the heavy atom sites used to calculate them. Moreover, the changes were large so that the number of cross-over terms might not be negligible. In fact, with the TCA derivative this number was some 20% of the total data. Coupled to this was the fact that only reflections with h+1 = 2n were present making the number of data available for the difference Fouriers about 50 unique terms per projection. Also, the low order terms are liable to large changes on account of the change in mother liquor when the heavy atom is added: this could have caused certain spurious effects although no one term of those used seemed to dominate the map causing a noticeable pattern throughout. Finally, the change in cell dimensions meant that the difference between native and derivative would not only be restricted to the extra heavy atoms but would probably involve some changes in the protein as well.

All of these factors, especially crossovers and the paucity of data for which phases were avilable, led to a very inconclusive result from the two-dimensional work. It was decided, therefore, that the three-dimensional, 6Å, difference Fourier should be obtained in the hope that this would clarify the situation.

SECTION II

SOLUTION STUDIES.

Materials.

The materials used in the following section were the following. The phosphates of sodium, potassium and ammonium, sodium chloride, A.S., potassium cyanide and potassium hydroxide were all Analar grade supplied by British Drug Houses Ltd., Poole, Dorset. Sodium tetrachloroaurate dihydrate was supplied also by BDH and was guaranteed greater than 98.5% pure. β-lactoglobulin AB was from Pentex Inc., Kankakee, Illinois and was recrystallised, freeze-dried and stored as an amorphous powder at $3^{o}C$ with a strip of paraffin-wax film sealing the lid to prevent moisture getting in. n-Butyl mercaptan and β-mercaptoethanol were both from BDH. The potassium gold cyanide was from Hopkin and Williams Ltd., Chadwell Heath, Essex.

Instruments

Ultraviolet and visible spectra were measured on either the manual recording Unicam SP500 or the automatic Unicam SP800 spectrophotometers. pH was measured, as for the aldolase work, with the EIL Model 23A Direct Reading pH meter equipped with a Jena micro-dual electrode as before. Polarimetric measurements were made on a Perkin-Elmer Model 141 Direct Readout Polarimeter using both the sodium lamp at 589 nm and the mercury vapour lamp at 436 nm. The polarimeter cells were 9.999 $\pm$ 0.001cms in length and were jacketed for constant temperature control from an external water bath. Weighings were made on a Stanton Model AD 2 beam balance with an accuracy (claimed)of 0.0002 g.

Precession photographs were taken on either a Supper Model D or a Jarrell-Ash 80-011 Buerger Precession camera with an 0.5 mm. collimator and a crystal to film distance of 7.5 cm. The film used was Ilford Industrial G and it was measured on a Joyce-Loebl Mk III Double-Beam Automatic Recording Microdensitometer with a lever-arm ration of 5:1. The X-ray generator was a Phillips P1011 fitted with a Phillips Fine-Focus tube and run at 40kV and 18mA.

The diffractometer used was the Hilger and Watts Y190 Linear Diffractometer again with a Phillips Fine-Focus tube run at 40kV and 20mA.

Note. Throughout this work the molecular weight of β-lactoglobulin was taken to be that of a single subunit, i.e. 18000.

Preliminary Experiment.

As has been stated above, studies in solution can be conveniently carried out using a polarimeter and a spectrophotometer. In order to see if TCA had any effect on the conformation and u/v spectrum, a short experiment was performed. This consisted of setting up flasks containing 1 microM β-lactoglobulin (M.W. 18,000) which were 25 microM in TCA and allowing them to equilibrate for eight hours before measuring the optical rotation and u/v spectrum. A similar preparation but without the TCA was made up at the same time as a reference. The pH was 7.51.

The value of the optical rotation, measured as $[\alpha]_D^{18}$, had increased from -38.0° to -55.6°. Also, the spectrum had changed quite markedly below 260nm. At first, it was thought that the changes were caused by a slow denaturation of the type found by

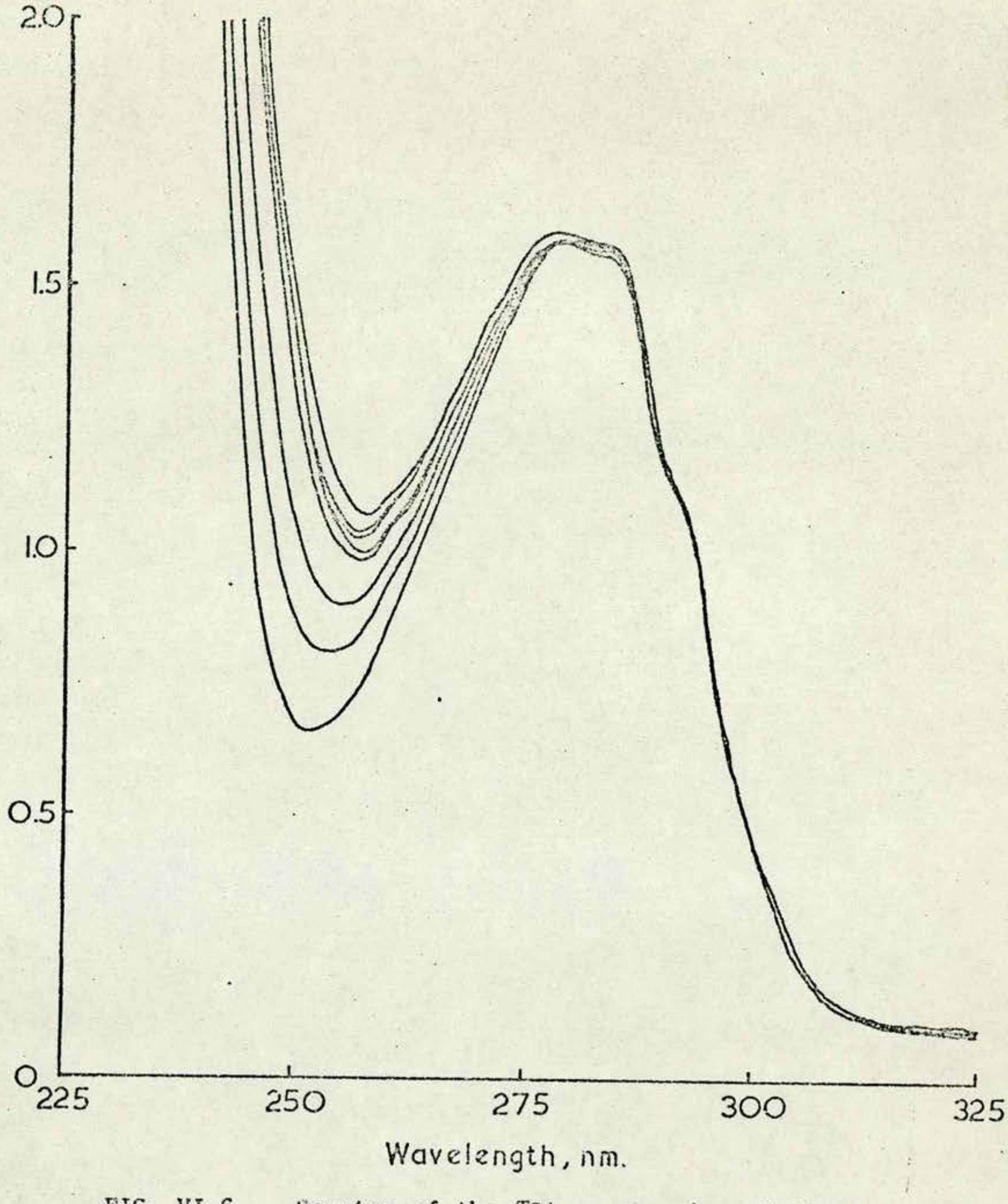
Pantaloni (1965), caught, as it were, midway between the R and S states but which, in time, would reach the final value of -74^{o}, characteristic of the unfolded, denatured molecule (Groves, Hipp and McMeekin (1951)).

However, measurements up to 72 hrs. after the original addition showed that both u/v and optical rotation results were unchanged. Thus, it appeared that a stable intermediate conformation might exist between the R and S states which could be examined both polarimetrically and spectroscopically. It was also noted that the quantity of TCA used in the solution did not have any appreciable absorption above 240nm.

Determination of the Number of Moles Reacting.

Before proceeding further it was necessary to characterise the reaction as well as possible by determining the number of moles of TCA per mole of protein required to cause the change, how stable a complex was being formed and which residue or residues were involved in the binding.

A stock 1.113×10^{-4}M solution of β-lactoglobulin was made up in 0.05M mixed phosphate buffer, pH7.6, and allowed to equilibrate at 3^{o}C overnight. Measurement of the pH of this solution at room temperature then showed that the slight buffering action of the protein had caused it to change to 7.52. Also, a stock 0.0249M TCA solution was made up by weighing out the $NaAu(Cl)_4.2H_2O$, then slightly more than four times the equivalent of KCN and dissolving them both in a few mls. of 0.05M phosphate buffer, pH 7.6, bringing the pH carefully back down to 7.8 with dropwise addition of 2M HCl



<u>FIG. VI.6</u> Spectra of the TCA complex in solution recorded on the SP800. The pH was 7.6, the protein was 1.02×10^{-4} M and varying amounts of 0.0249M TCA were used. x = 0.01, 0.02, 0.04, 0.06, 0.08, 0.10 mls.

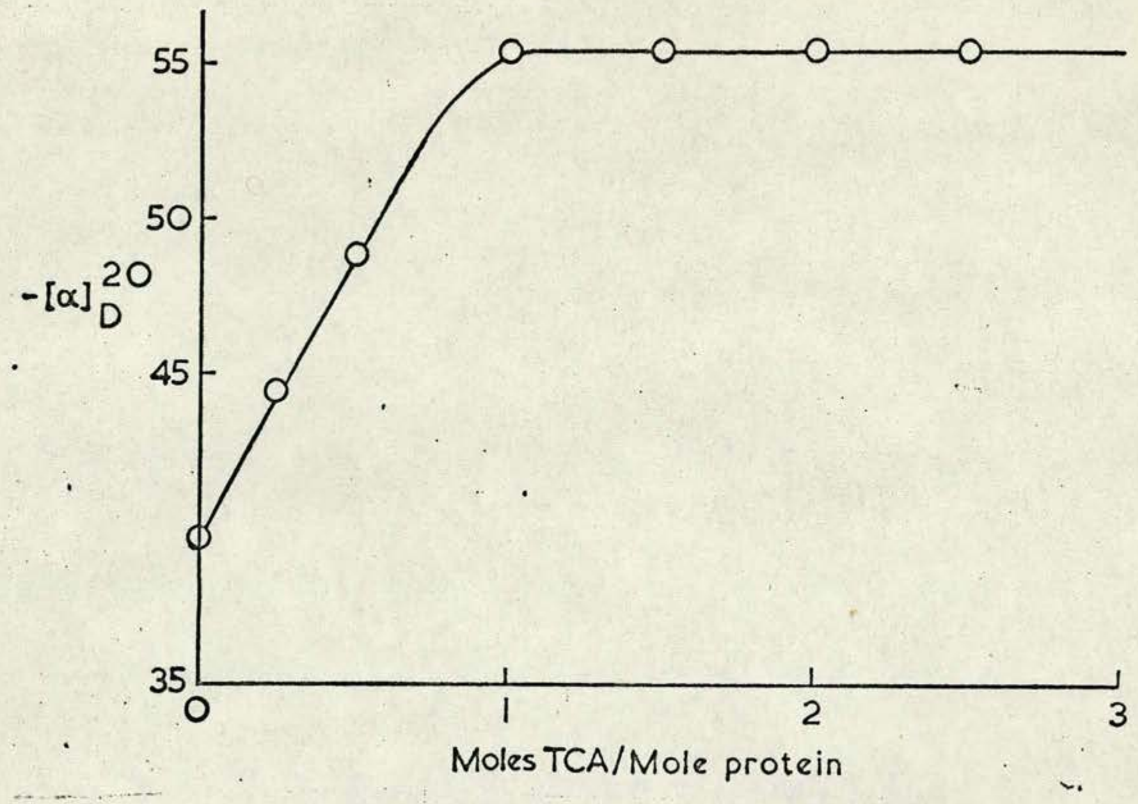


FIG. VI.7 The specific rotation of the solutions from Fig. VI.6 at 18°C and 589nm.

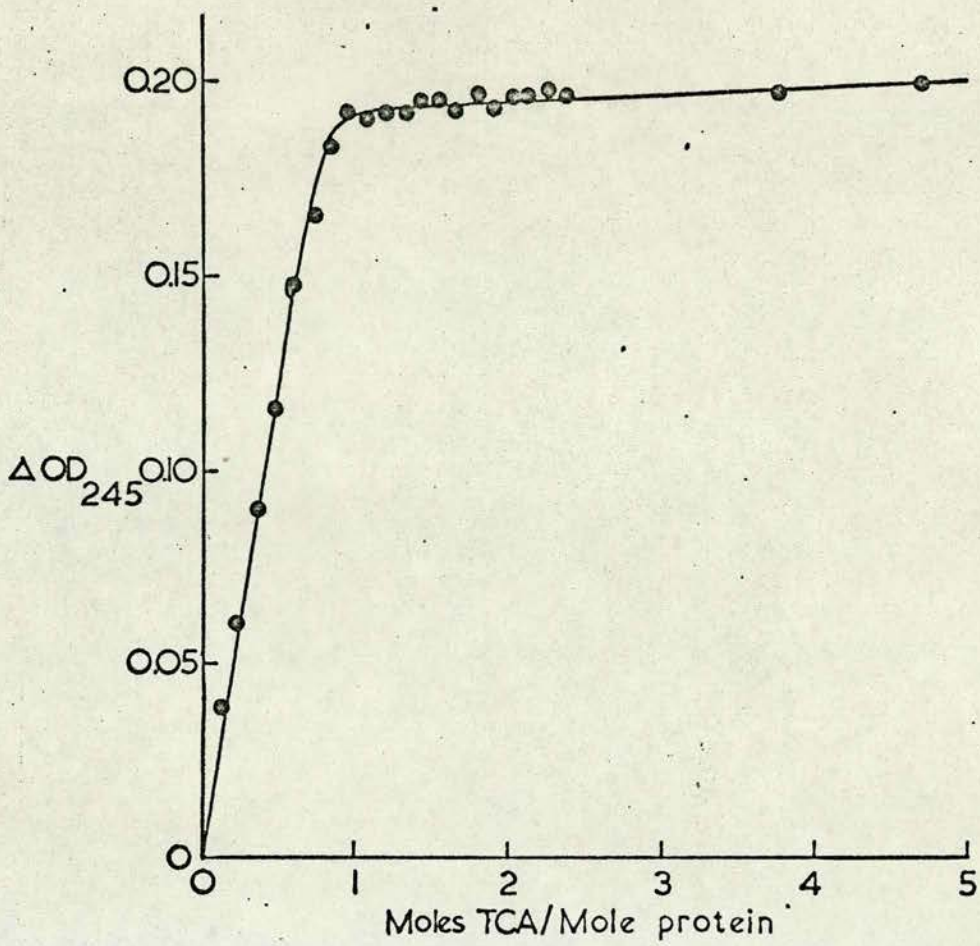


FIG. VI.8 The stoichiometry of the complex at pH 7.6 in 0.1M phosphate as measured on the SP500. The protein was 2.01×10^{-5}M.

and then making the solution up to the mark with more buffer. Preliminary spectra measured with the SP500 showed that, at the concentrations used, neither CN^- or Cl^- ions caused any detectable alteration in the protein spectrum.

A series of flasks was then set up containing:

9.00mls protein solution

x mls TCA solution

(1-x)mls 0.05M phosphate buffer.

and allowed to equilibrate overnight at 3^oC. The optical rotation was then measured at 18^o and 589nm (Na D line). The u/v spectra were run on a Unicam SP800 recording spectrophotometer, using the fast scanning speed. These spectra are shown in Fig. VI. 6. In Fig. VI. 7 the optical rotation data are plotted.

In order to check the spectrophotometric data on the more accurate SP500 and in order to use it in its optimum range, $(0.1 \leq E \leq 0.4)$ a fresh series of solutions were made up as follows:

1.00mls. 1.002×10^{-4}M protein in 0.1M phosphate

x mls. 2.50×10^{-4}M TCA

(4-x)mls. 0.1M phosphate buffer, pH 7.6.

The optical densities were measured at 245,255 and 280nm. after 36 hrs. The value of E_{280} was used as a check on the protein concentration, the mean value being taken as the true one and all others scaled to it. The data for 245nm are shown in Fig. VI. 8.

Difference spectra with β-lactoglobulin as the reference showed there to be what appeared to be a maximum at 240nm. However, this was found to be the SP800's failing to function properly because the reference solution was becoming opaque at that wavelength. Fivefold diluted samples were measured manually on the SP500 and these data

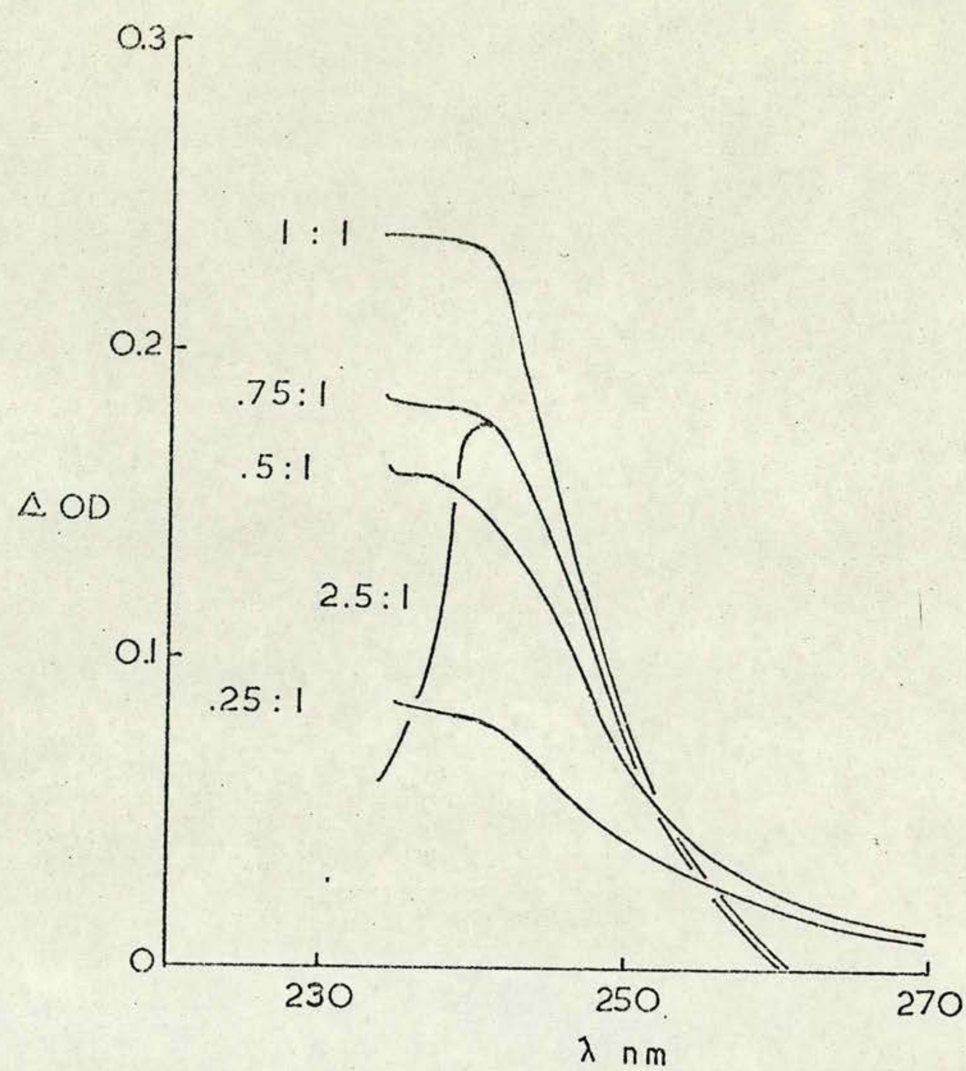


FIG. VI.9 Difference spectra of the samples from Fig. VI.6 diluted fivefold and recorded on the SP500.

are shown in Fig. VI. 9. It can be seen that, at ligand : protein ratios greater than 1 : 1, a peak, or possibly just a shoulder, develops at 242nm. It is still, however, close to a very large peak which requires rather larger slit widths than normal and this was thought to be a source of inaccuracy. Further dilution to concentrations where this was no longer a problem caused the peaks at 280nm and 242nm to be almost indistinguishable from the 'background' on the side of the main peak.

These results all indicated that there was a stoichiometric reaction between β-lactoglobulin and TCA resulting in the formation of a 1 : 1 complex. With a large excess of TCA there was a slight indication of a further increase in the optical density but this did not seem to be nearly such a definite reaction.

pH Dependence of The Reaction

With what was the TCA forming a complex? Dunnhill and Green (1965) had shown that the rate of reaction with pCMBS was affected by the pH in such a way that below the Tanford transition the reaction was very slow whilst above it, it proceeded with great speed. pCMBS however, is a reagent which is specific for sulphydryl groups. Because of the lack of further evidence as to the group or groups involved in binding the TCA, it was decided to examine the effect of pH on the reaction between β-lactoglobulin and the gold complex. An effect similar to that of Dunnhill would indicate that either the sulphydryl group itself or the region of protein in which it was situated, was involved.

Once more a stock protein solution was made up in 0.1 M

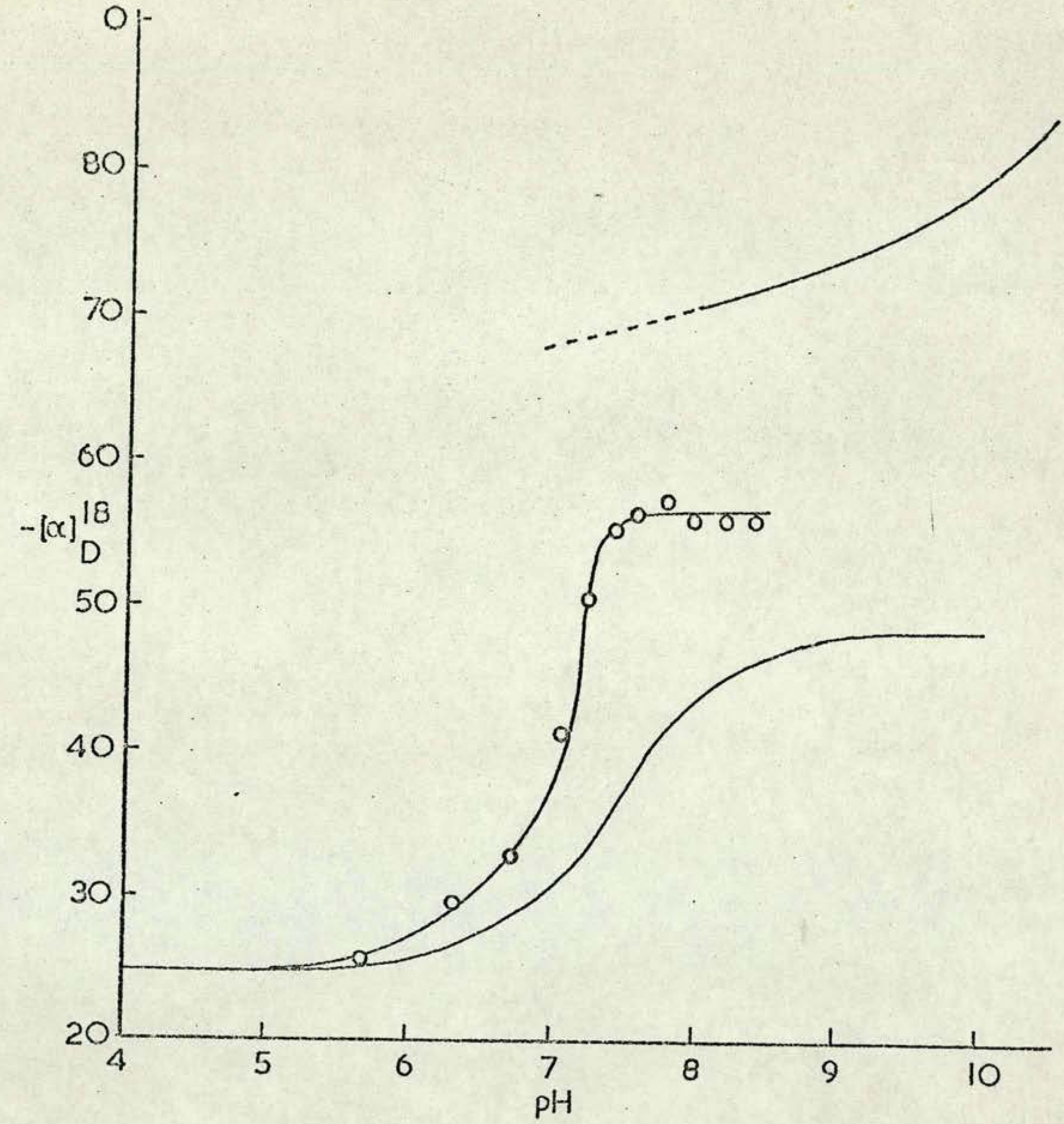


FIG. VI.10a pH dependence of the formation of the TCA-protein complex depicted as a change in specific rotation at 589nm. The normal change as obtained by Tanford is also shown along with the data of McMeekin for the denatured protein. The protein was 0.999×10^{-4} M, the TCA was 2.5×10^{-4} M and the temperature was 18^{o}C.

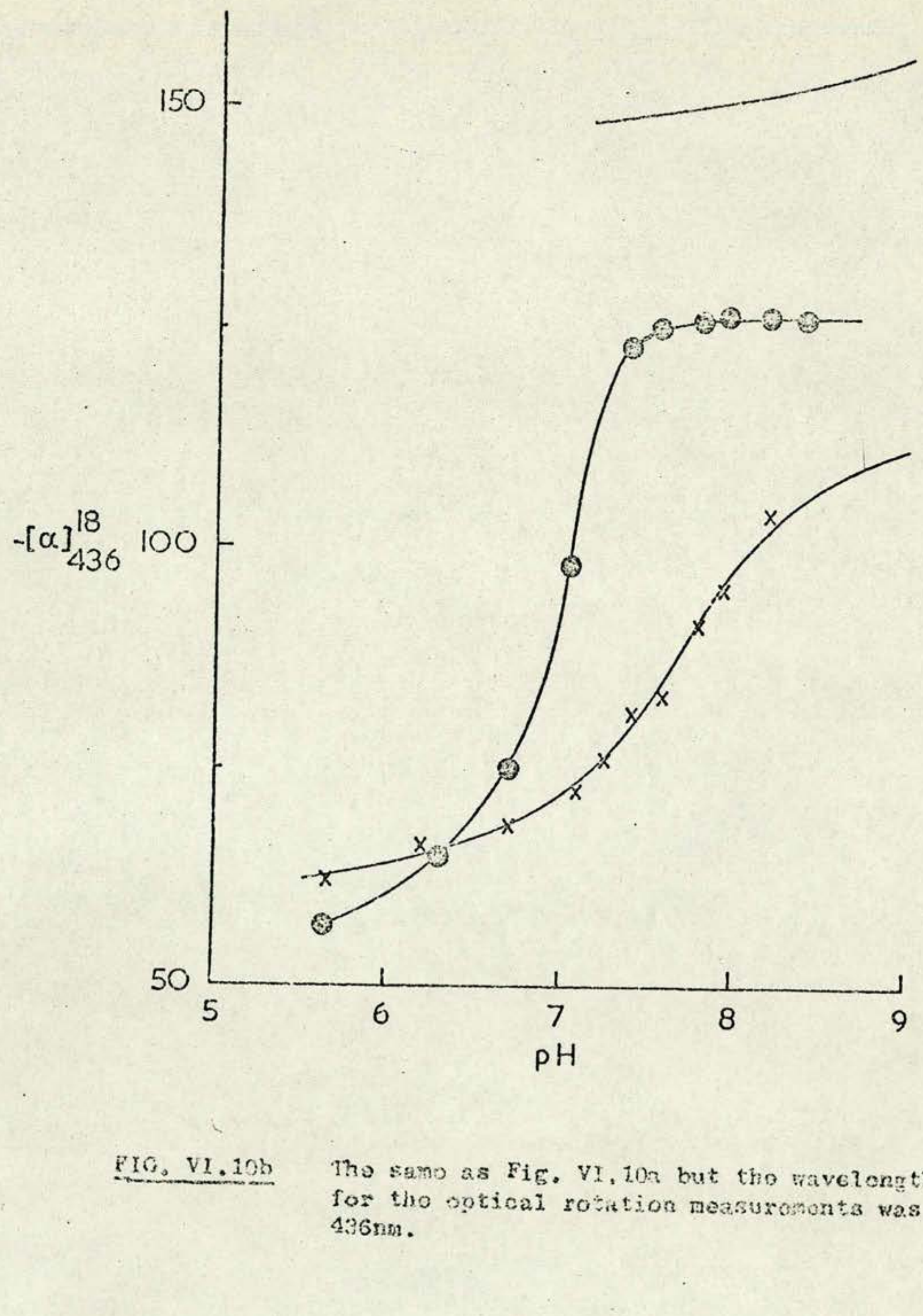


FIG. VI.10b The same as Fig. VI.10a but the wavelength for the optical rotation measurements was 436nm.

phosphate buffer at pH 5.5. This time, however, instead of allowing the solution to equilibrate, the flask was made up as shown below:

5.00 mls. 1.997×10^{-4} M protein solution

1.00 mls. 0.0025M TCA solution

4.00 mls. water

This gave a molar ratio of TCA : protein (M.W. 18,000) of 2.5 : 1, sufficient to ensure complete reaction. A reference solution was made up in an identical fashion save that the TCA was replaced by water. The stock protein solution was then titrated to about pH 6.0 with 4M KOH and a flask made up as above along with another reference. This complete sequence was repeated, raising the pH in steps of about 0.3 pH units, until pH 8.5 was reached. The solutions were then set aside at 3°C to equilibrate overnight. Storage in the refrigerator was considered necessary in order to reduce the risk of denaturation and bacterial attack even although the rate of the reaction might be reduced.

Twenty four hours later, the solutions were allowed to warm to room temperature and measurements were made of the pH, optical rotation both at 589nm and at 436nm, and the u/v spectra (SP800). The optical rotation data are plotted as a function of pH in Fig. VI. 10 a and b. The reference data are also plotted along with the curve obtained by Tanford, Bunville and Nozaki (1959). As can be seen from these figures, the reaction, like that of pCMB, is pH dependent. Unlike the pCMB, however, a stable conformation midway between the R and S states has appeared on the binding of one TCA molecule per protein subunit.

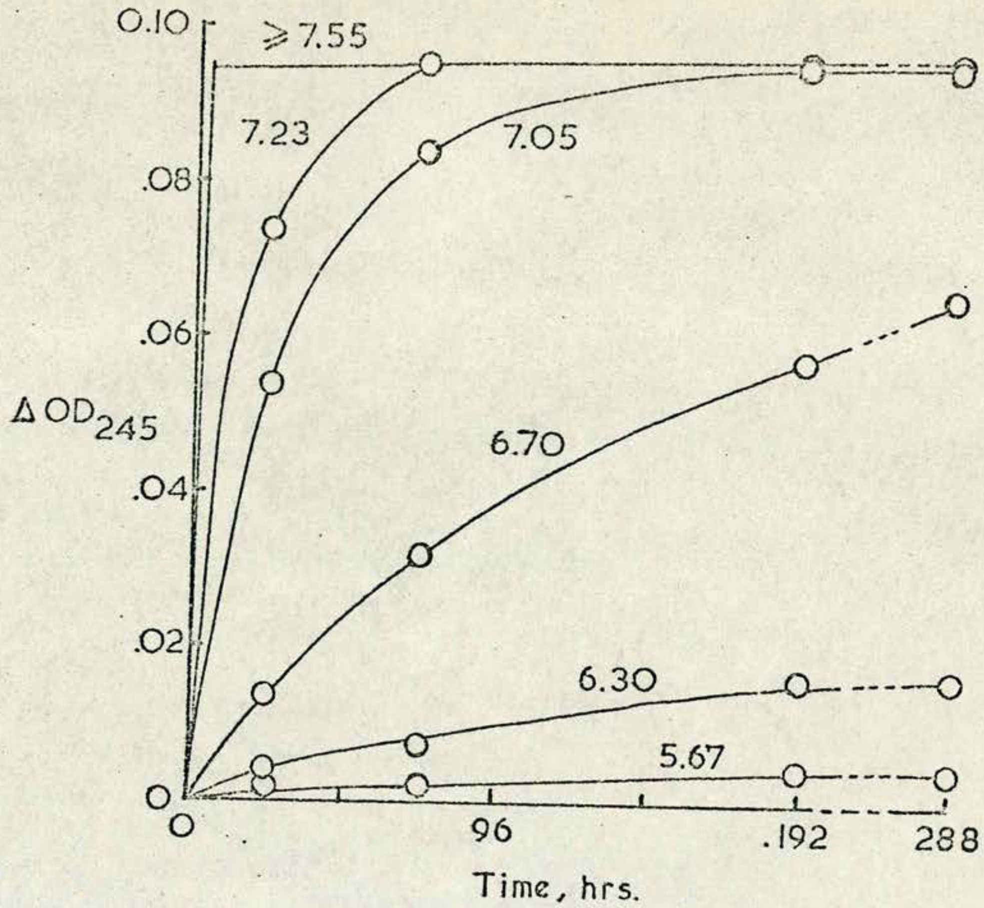


FIG. VI.11 Change of O.D. at 245nm as a function of time for the samples in Fig. VI.10.

The u/v spectra, as for the earlier experiment,parallel the optical rotation results. As before, the absorption became too great below 240nm so that no useful measurement was obtained below this value. It was also found that the value of the optical density increased slightly with time over a period of days for those samples which had not undergone the full reaction. Most change was noticed for the samples close to the pH value at which full reaction could occur, as though the reacted conformation was the more stable. This can be seen in Fig. VI. 11 where the data are plotted for the various pH values. The difference data were obtained from the mean value of the optical density measured at 245nm for all of the reference solutions and from the optical density at 280nm for both reference and sample solutions, this latter value being used to put all of the data on the same scale. Since these data were measured over a period of ten days and since only four readings for each pH were taken in this time, it is not inconceivable that the errors shown are too small. However, the changes near pH 7.2 were large enough to be significant and indicated that a change in the conformation could be induced at pH values lower than normal.

As there were several of the flasks from the above experiment which were fully reacted, it was thought worthwhile to see if the reaction was reversible by lowering the pH of each of the flasks by about 1.5 pH units, by the cautious addition of 2M HCl. Some slight precipitation of denatured protein had occurred overnight during equilibration, possibly started by the uneven mixing of the acid. This was removed by filtration through Whatmans' No. 1 filter paper before the optical rotation measurement. These results are

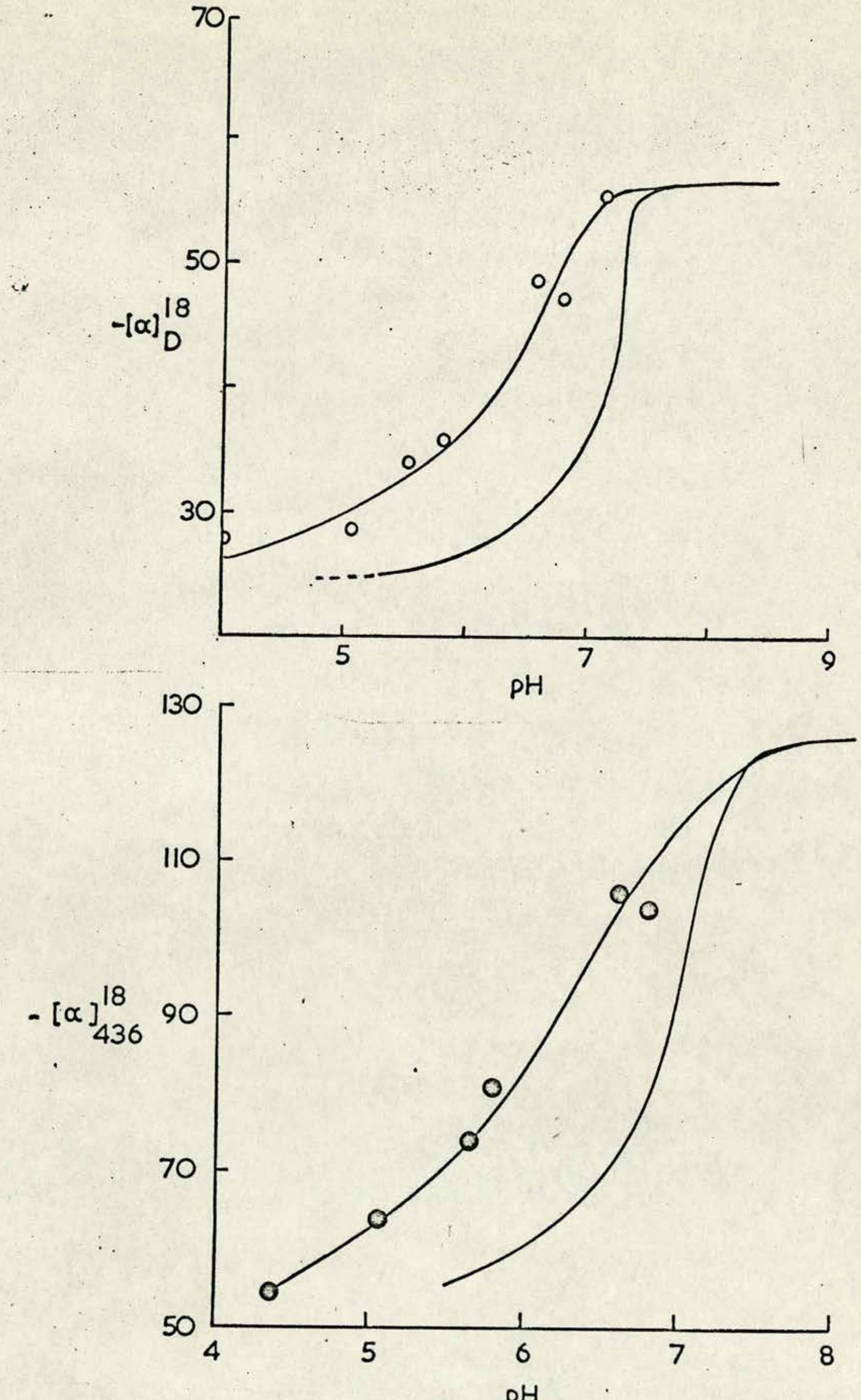


FIG. VI.12 Optical rotation measurements showing the back reaction of the TCA-protein complex using some of the samples from Fig. VI.10.

shown in Fig. VI. 12 together with the normal curve for comparison. It was clear from this that the TCA was not being bound irreversibly but it was not clear if the ligand-protein "bond" was being broken as a result of strong interactions with some of the protein side-chains which were being forced back into their low pH positions by the rest of the molecule or if the slightly acid solution was enough to hydrolyse the "bond" between TCA and protein, thereby allowing the protein to take up its N state.

Determination of the Group of Attachment.

As mentioned above, the fact that the reaction with TCA was pH dependent could mean one of two things: either the TCA was binding to the free sulphydryl group or it was binding to some other group closely associated to the SH in the tertiary structure. Because of the ability of thiols to reduce $AuCl_4^-$ to colloidal gold, it was felt worth investigating the reaction of TCA with some small thiols in order to see if a parallel reaction could be obtained to that occurring with β-lactoglobulin. In other words, could sulphydryl groups reduce TCA and, if so, was there a characteristic absorption band around 240nm?

Latimer (1952) gives the oxidation potentials of gold ions as shown below.

$$\begin{tikzcd} Au^{3+} \arrow[r, "1.41v"] \arrow[rr, bend right, "1.50v"'] & Au^{+} \arrow[r, "1.68v"] & Au & AuCl_4^- \arrow[l, "1.00v"'] \end{tikzcd}$$

Also, values for the stability constants of $Au(CN)_2^-$ and TCA obtained from the Chemical Society Special Publication No. 17 (Stability

Constants) are

$$Au^{3+} + 4CN^- \rightleftharpoons Au(CN)_4^- \quad , \quad \log\beta_4 \simeq 56$$

$$Au^{+} + 2CN^- \rightleftharpoons Au(CN)_2^- \quad , \quad \log\beta_2 \simeq 38$$

Using the above data in a Nernst-type equation (see, for example, Vogel (1962))

$$E = E_o + RT/nF.\ln\,[\text{Oxidant}]/[\text{Reductant}]$$

where E is the required oxidation potential, E_o is the standard potential of the material in question, n is the number of electrons involved in the oxidation/reduction, R is the gas constant, T is the absolute temperature, F is the Faraday and the square brackets refer to the concentrations of the oxidant and the reductant, it is possible to determine the oxidation potential of the reduction of TCA to $Au(CN)_2^-$. The values obtained are shown below.

$$E_{TCA\text{-}Au} = 1.50 + 0.059/3.(-56) = 0.40v.$$

$$E_{Au(CN)_2^-\text{-}Au} = 1.68 + 0.059/1.(-38) = -0.56v.$$

$$E_{TCA\text{-}Au(CN)_2^-} = 1.41 + 0.059/2.(38-56) = 0.88v.$$

These form the consistent set shown below

$$\begin{tikzcd} Au(CN)_4^- \arrow[r, "0.88v"] \arrow[rr, bend right, "0.40v."'] & Au(CN)_2^- \arrow[r, "-0.56v."] & Au \end{tikzcd}$$

The value obtained is somewhat higher, for the reduction of TCA to $Au(CN)_2^-$, than that reported by Remick (1947), E = 0.5v in dilute H_2SO_4. His value is not determined directly, being estimated by comparison with other redox reactions. It must also be admitted that the value obtained above depends on the values given for the stability constants for the complexes and these could be out by up to 10% .

The oxidation of a sulphydryl to a disulphide can be written as

$$2RSH \rightleftharpoons R\text{-}S\text{-}S\text{-}R \ + \ 2H^{+} \ + \ 2e^{-}$$

The direct measurement of the redox potential of this reaction is made difficult by the affinity of thiols for mercury, usually present as an electrode. Values for various thiols have been given by Calvin (1954) and, more recently, for the cystine/cysteine couple by Gorin and Doughty (1968) and Jocelyn (1967). Eldjarn and Pihl (1957) compared several thiols with the glutathione reaction and found them all to be within 0.1v of the reference. It would appear therefore that the value for the glutathione redox reaction may be regarded as a fair approximation to the value of E for a thiol. For glutathione, Scott, Duncan and Ekstrand (1963) have determined enzymatically a value of E = 0.17v for the half-reaction

$$\tfrac{1}{2}GSSG \ + \ H^{+} \ + \ e^{-} \rightleftharpoons GSH$$

This meant that the oxidation of a thiol was quite possible by TCA.

It was therefore possible that the reaction of the TCA molecule was with the free sulphydryl group in β-lactoglobulin. But it was still necessary to show that the spectral change was caused by this interaction. To achieve this, TCA was reacted with β-mercaptoethanol (ME) in various ratios, there being the same quantity, about 5 micromoles, of ME present per determination. The spectra of the ME solution alone and then with TCA are shown in Fig. VI. 13, the solutions having been diluted tenfold before measurement (SP500). The solutions were also measured immediately after being

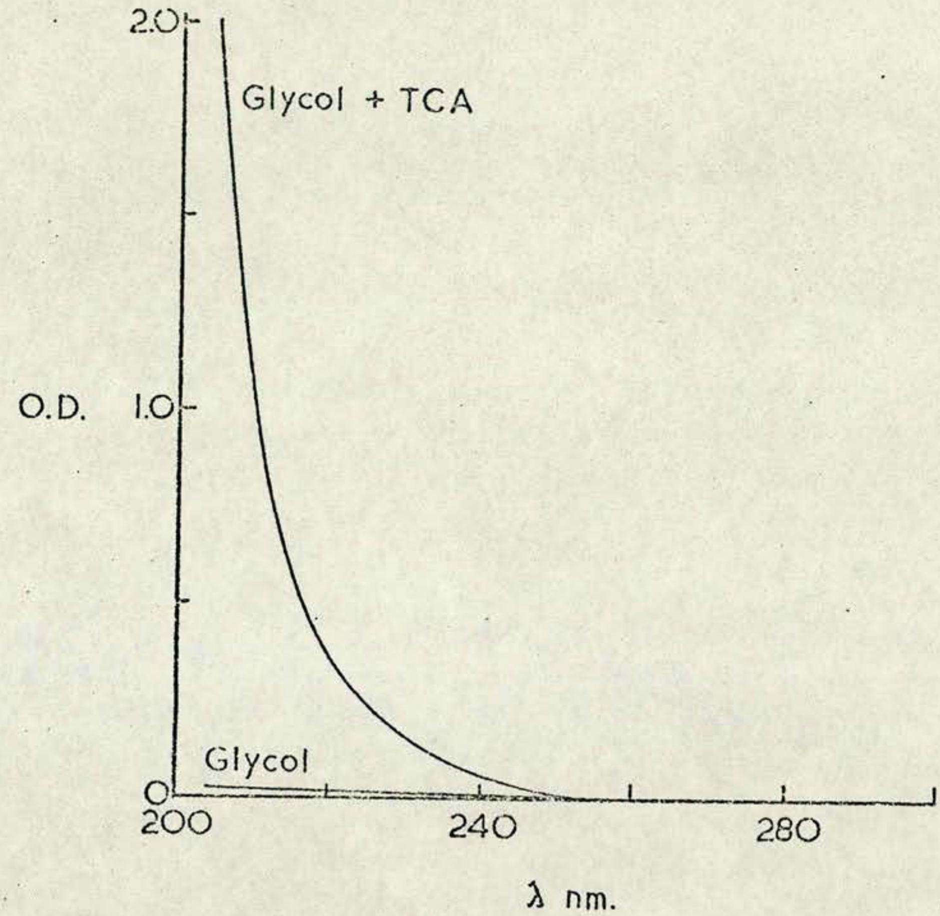
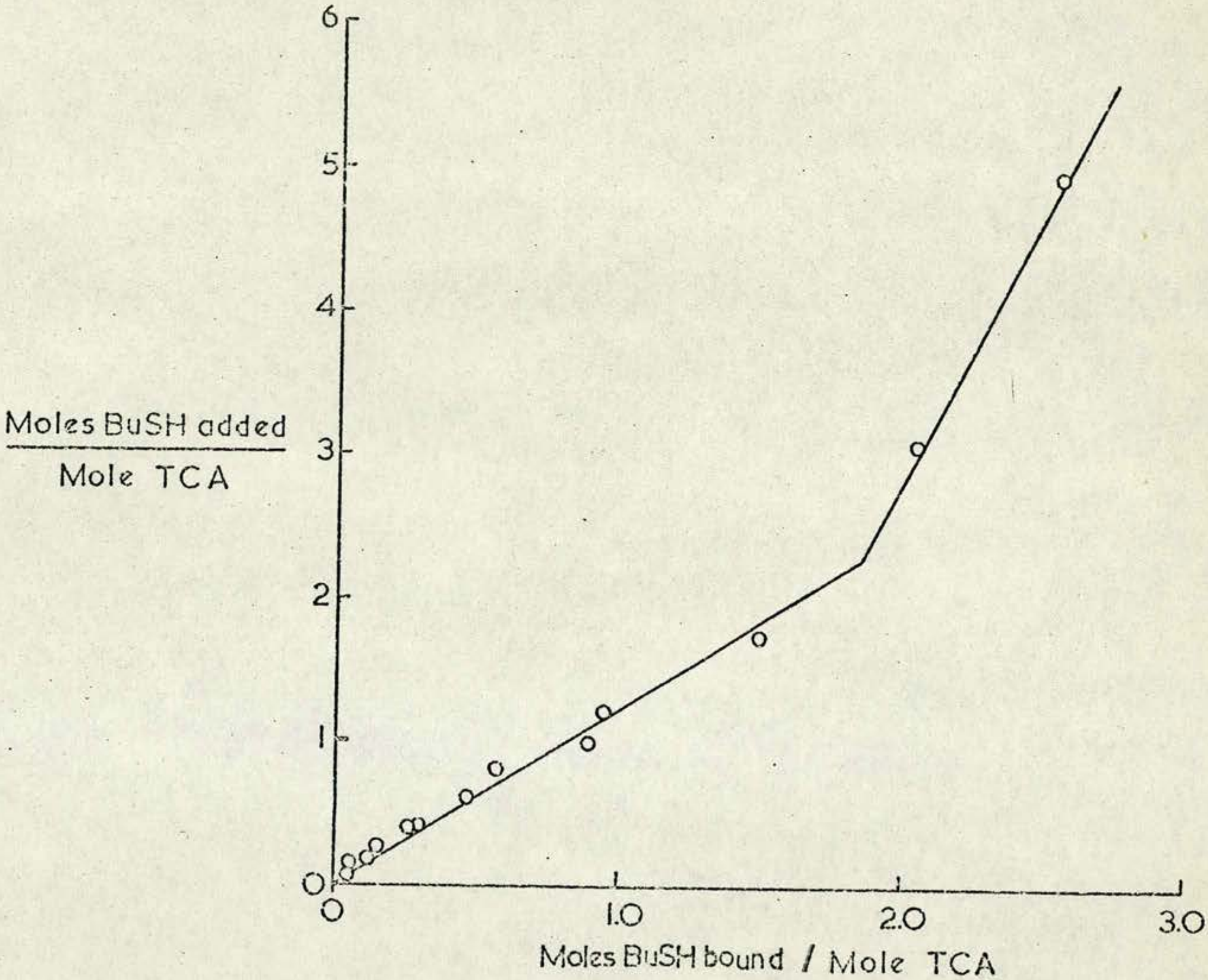


FIG. VI.11 Spectra of 0.02 mls glycol in 24.98 mls water, and 4.8 mls of this solution with 0.2 mls 0.025M TCA. The pH was 7.2.



<u>FIG. VI.15</u> Back titration with DTNB of the free sulphydryl remaining in solution after reaction of a fixed amount of BuSH with varying amounts of TCA.

made up because it was found that they changed with time. The figure shows the development of peaks at 291nm and a broader one between 230nm and 240nm. This seemed to indicate that the peak around 240nm was caused by the interaction of an SH group with Au(III), and pointed to the TCA's reacting with the sulphydryl group in the protein.

In order to check that it was the thiol rather than the hydroxyl group in ME which was responsible for the formation of the peaks observed at 290nm and 240nm, the spectrum of a solution of ethylene glycol of equal molarity to that of the ME was measured both with and without the same quantity of TCA as was used to give the more intense peak in Fig. VI. 13. These spectra are shown in Fig. VI. 14 where it can be seen that there are no peaks corresponding to those in ME above 220nm.

As a final check that the sulphydryl group alone was responsible for the peaks, reaction was carried out with n-butyl mercaptan (BuSH). This compound is sparingly soluble in water, but a saturated solution was found to be about 6mM in SH, by use of 5,5'-dithio-bis-(2-nitrobenzoic acid), DTNB (Ellman (1959)). Unfortunately, however, it was found that the BuSH evaporated readily thus making quantitative measurements of thiol concentration inaccurate. Consequently, treating constant quantities of BuSH with varying quantities of TCA and monitoring the reaction by assaying the amount of free SH in solution, although theoretically possible, in practice turned out to be rather unconvincing, probably made worse by the intermediate mercaptide of TCA being unstable. These data are shown in Fig. VI. 15 where it can be seen that there is a change of slope at about 2 moles of BuSH per mole of

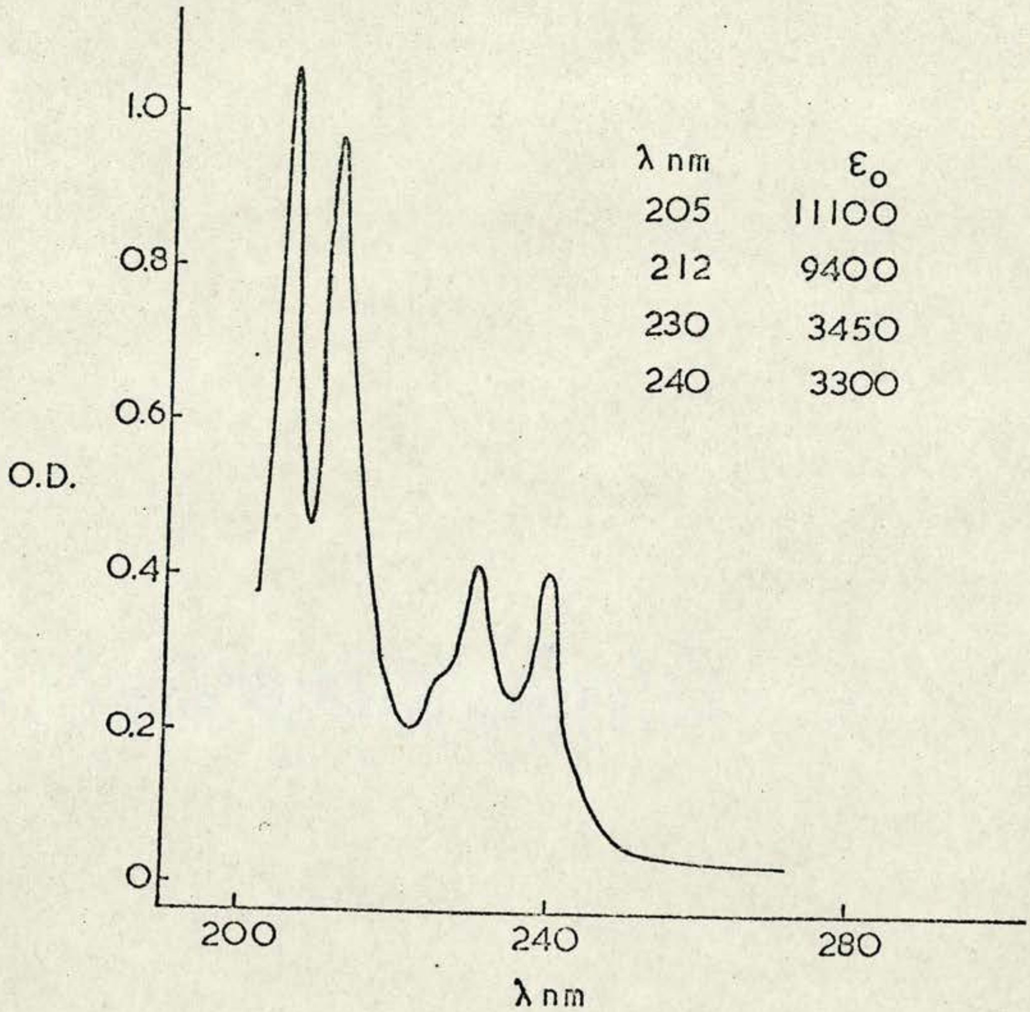


FIG. VI.16a Spectrum of a 9 x 10^{-5}M solution of $KAu(CN)_2$.

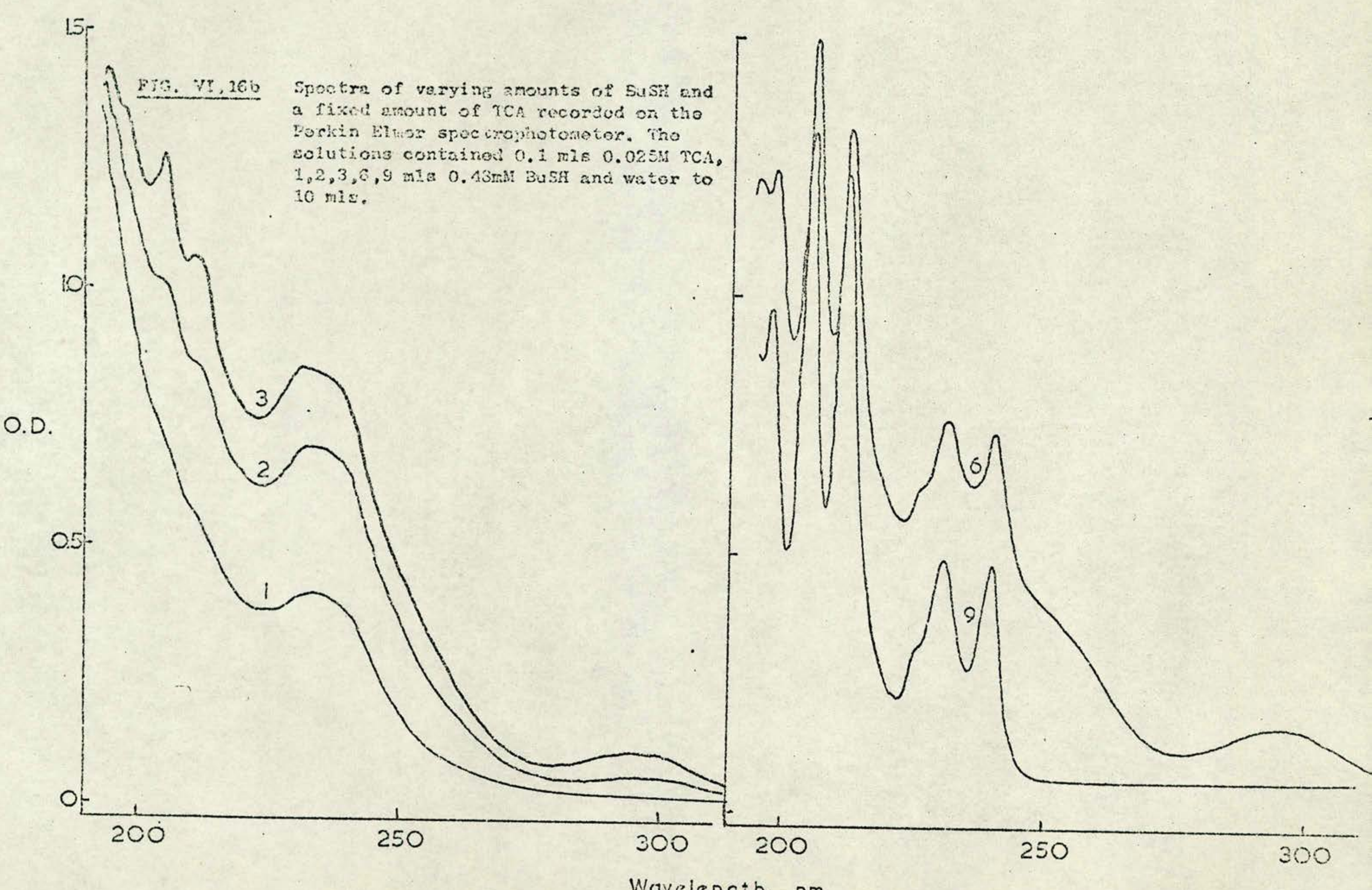


FIG. VI.16b Spectra of varying amounts of BuSH and a fixed amount of TCA recorded on the Perkin Elmer spectrophotometer. The solutions contained 0.1 mls 0.025M TCA, 1,2,3,6,9 mls 0.43mM BuSH and water to 10 mls.

TCA. This corresponds to the reaction

$$Au^{III}(CN)_4^- + BuSH \rightleftharpoons Au^{II}(CN)_3^- + HCN + BuS^{\bullet}$$

$$Au^{II}(CN)_3^- + BuSH \rightleftharpoons Au^{I}(CN)_2^- + HCN + BuS^{\bullet}$$

$$2BuS^{\bullet} \longrightarrow Bu\text{-}S\text{-}S\text{-}Bu$$

giving an overall reaction of

$$Au^{III}(CN)_4^- + 2BuSH \rightleftharpoons Au^{I}(CN)_2^- + 2HCN + BuS\text{-}SBu.$$

Whilst BuSH is sparingly soluble,the disulphide derived from its oxidation is insoluble. Therefore, in flasks containing greater than a 1:1 ratio of BuSH to TCA, an emulsion should form from the production of the insoluble disulphide. This was observed after the solutions had been allowed to stand overnight. It was also noticed, incidentally, that there was almost no smell of BuSH (a characteristic and unpleasant one) in flasks containing less than a 2:1 ratio.

Most convincing of all, however, were the u/v spectra of these same solutions measured on a Perkin-Elmer Model 402 recording spectrophotometer*, using an equal BuSH concentration as reference. Fig. VI. 16 (a) shows the spectrum of $KAu(CN)_2$ solution and Fig. VI. 16 (b) shows the spectra of increasing ratios of BuSH to TCA. At ratios approaching 2:1 there is evidence of the appearance of the $Au(CN)_2^-$ spectrum,which is as expected. Above 2:1, there is no peak at 290nm showing that this peak arises from the intermediate formed in the reduction of TCA by thiols.

On page 89, it was stated that $AuCl_4^-$ was discarded because of

---

*Kindly lent on approval by Perkin-Elmer Ltd., Beaconsfield,Bucks.

its ease of reduction to Au. Evidence for this came not only from its oxidation potential but also from an observation of the formation of a precipitate produced on treatment of a solution of tetrachloroaurate with ME. However, filtration of the solution through a Millipore filter RAWP 047 00 with a pore size of 1.2 microns, yielded a clear solution with a shoulder in its spectrum at about 240nm. This was further evidence that the band at this wavelength is associated with the reaction between Au(III) and a thiol.

Therefore, all of the evidence pointed to the two bands at 240nm and 290nm being a direct result of the binding of a sulphydryl group to trivalent gold. A search through the literature up to 1969 has not revealed any u/v spectra of compounds containing a single Au^{III}-S bond. However, the compound dibromo-N,N-di-n-butyl dithiocarbamatoaurate (III) has a spectrum which has shoulders at the appropriate wavelengths, although the interpretation has not been put forward by the authors (Buerskens, Cras and Steggerda (1968)) since there are several peaks running together between 220nm and 350nm.

<u>Proposed Reaction and the Complex Formed with the Protein.</u>

Thus, the reaction of TCA with β-lactoglobulin, like that of pCMBS, involves the free cysteine and is dependent on the conformations being such that the sulphydryl group is available. This allows a reaction to be postulated:

$$\text{Protein-SH} + Au(CN)_4^- \rightleftharpoons \text{Protein-S-}Au(CN)_3^- + HCN$$

Here the reaction must stop unless some means of supplying a further electron for the reduction of the Au, can be found from the solution or, possibly, the protein. Similar replacement reactions have been examined by Cattalini, Marangoni and Martelli (1968) using dibenzyl thioether (Bz_2S), according to the reaction

$$AuCl_3am \quad + \quad Bz_2S \rightleftharpoons AuCl_3Bz_2S \quad + \quad am$$

(am: heterocyclic nitrogen base)

By studying the reaction kinetics of this system in acetone, where the reduction of the gold by the excess Bz_2S is kept to a minimum, Cattalini and his coworkers have been able to show that the reduction of Au(III) to Au(I) by the thioether is a two-stage process, the first being the replacement reaction followed by the second, slower, reduction step. Some earlier work in the same laboratory (Cattalini, Orio and Tobe (1967)) using $Au(CN)_2Cl.am$ instead of the trichloro compound, showed that the cyano group had little effect on the replacement stage.

Therefore, because of the unavailability of another sulphydryl group and because of the very characteristic $Au(CN)_2^-$ spectrum which would have shown up in the diluted difference spectra had it been formed, the postulated reaction was taken to be the explanation of the various experimental observations. Extra stability for the complex could be gained from the formation of hydrogen bonds between the remaining cyanide groups and nearby amino or hydroxyl groups in the protein.

A small piece of extra evidence in support of the above findings was obtained by measuring the optical rotation of 5:1 $Au(CN)_2^-$ to protein solution at pH 7.8 as for the TCA measurements. This showed there to be no change in the specific rotation from that of a reference sample made up with water in place of the gold complex.

Direct Evidence for the Postulated Complex.

Returning to the reduction of TCA by BuSH, and bearing in mind the requirement of substitution before reduction put forward by Cattalini et al. (1968), the complete reaction can be written as

$$Au^{III}(CN)_4^- + BuSH \longrightarrow Au^{III}(CN)_3\text{-}S.Bu^- + H^+CN^-$$
$$Au^{III}(CN)_3\text{-}S.Bu^- \longrightarrow Au^{II}(CN)_3^- + Bu.S^{\bullet}$$
$$Au^{II}(CN)_3^- + BuSH \longrightarrow Au^{I}(CN)_2^- + Bu.S^{\bullet} + H^+CN^-$$
$$Bu.S^{\bullet} + Bu.S^{\bullet} \longrightarrow Bu.S\text{-}S.Bu$$

Because of the instability both of the Au(II) and of the radical, $Bu.S^{\bullet}$, no reduction can occur unless the way is clear for a complete reaction. That is, unless there is some means of supplying another electron, the substituted Au(III) complex will be stable. Consequently, a direct proof of the existence of the TCA-protein complex should be obtained by supplying the second sulphydryl group thus allowing the formation of a disulphide in the protein and releasing $Au(CN)_2^-$ to the solution.

Obviously it would be better to try to detect the dicyanoaurite without the intense protein band's interference. By containing the β-lactoglobulin in a dialysis sac, it was thought that such a separation could be obtained. However, the Visking tubing, supplied by the Scientific Instrument Centre, 1, Leeke Street, London, contains about 0.1% sulphur (Manufacturer's specification). This was determined as equivalent to 0.7 micromoles of sulphur per centimetre of the 24/32" diameter tubing available in the laboratory. Of course, the sulphur content might not all be present

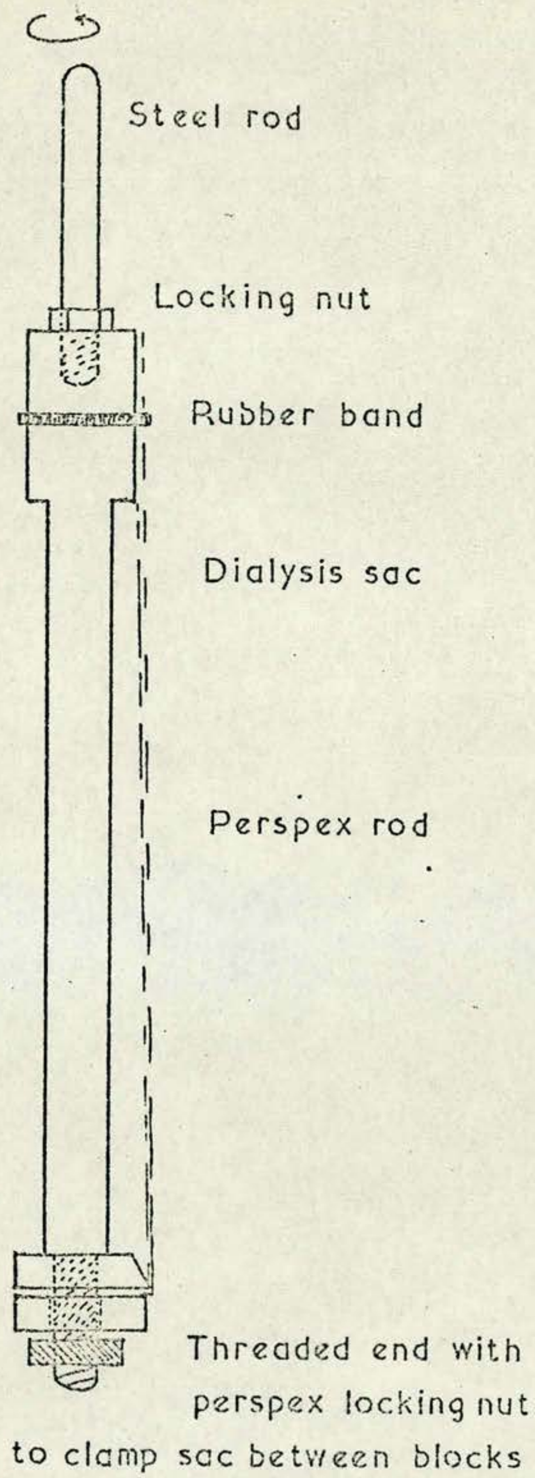


FIG. VI.17 Drawing (not to scale) of the device used for dialysis.

as SH, but even half of the quantity determined above present as the free thiol would be enough to interfere seriously with the results if the earlier, convenient concentrations were to be used. As a precaution, therefore, the tubing was boiled in an approximately 2% solution of $NaHCO_3$ for 5-10 minutes and then washed with copious quantities of water before being stored in water at 3^oC.

Before trying the addition of BuSH to the protein-TCA complex, an experiment was carried out to see how firmly the TCA was attached. The apparatus used is shown in Fig. VI. 17, the benefit of such a device being that equilibrium is achieved in five or six hours because the relative surface area is so large and the motion ensures even mixing of the dialysing solution. The method of holding the sac, too, is useful for access to the dialysate if samples of this have to be removed from time to time.

A stock 0.001M solution of β-lactoglobulin which was 0.0008M in TCA was prepared in 0.1M mixed phosphate buffer, giving a pH after equilibration of 7.5. The reaction was allowed to reach completion by standing the flask for 48 hours in the refrigerator. After this time, 5 mls. aliquots were dialysed against 500 mls. distilled water for 15 hours. The dialysing solutions were then concentrated by evaporation and the spectra of these solutions showed no evidence of peaks above 220nm. Measurement of the spectrum of tenfold diluted samples of the dialysate gave the same ratios of E_{280}/E_{255} as a similarly

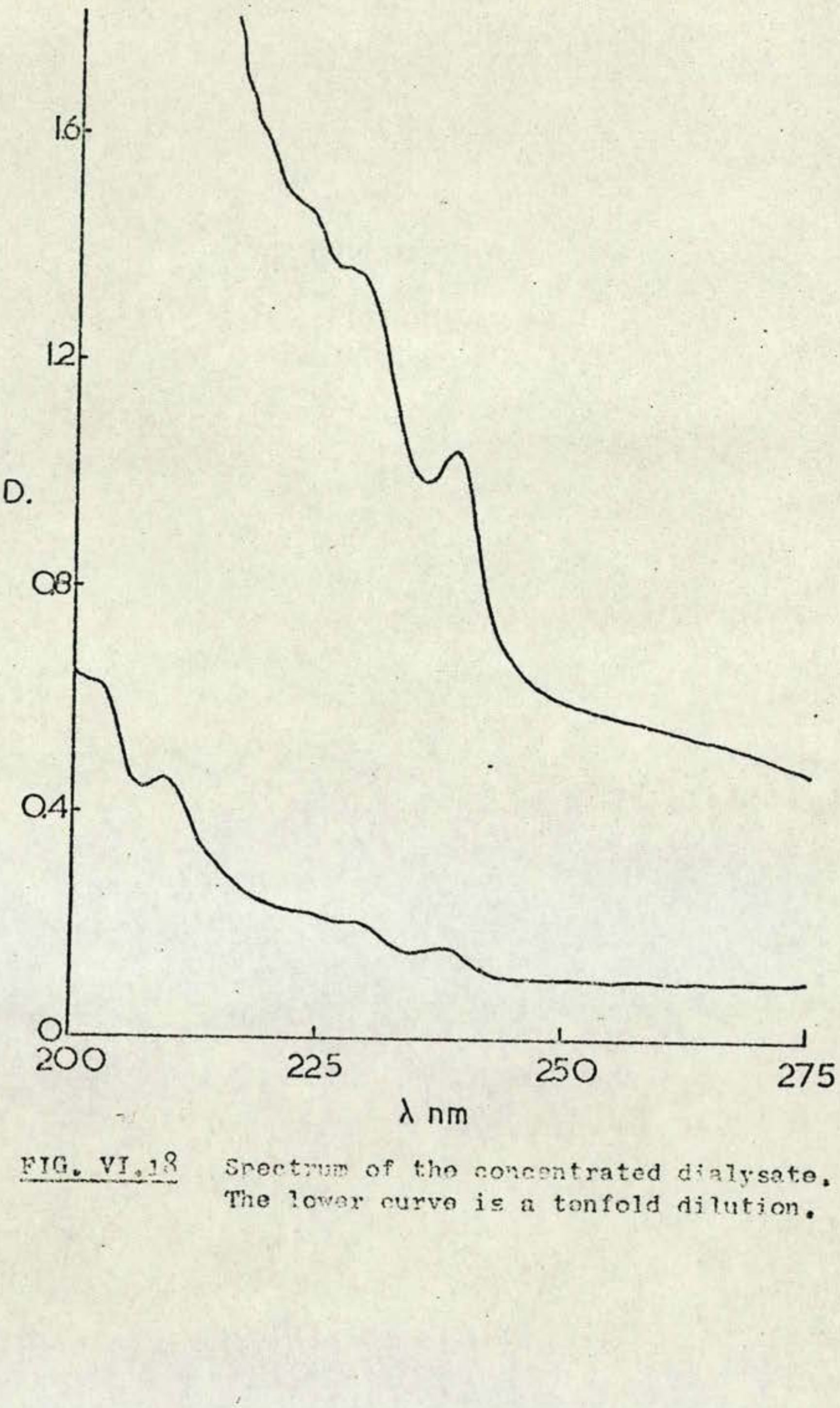


FIG. VI.18 Spectrum of the concentrated dialysate. The lower curve is a tenfold dilution.

diluted sample of the stock solution. These observations indicated that dialysis alone was not enough to remove significant amounts of the TCA, provided the length of time of dialysis was not increased. Use of a less than 1:1 ratio of TCA to protein ensured that there was a minimum of free TCA in solution.

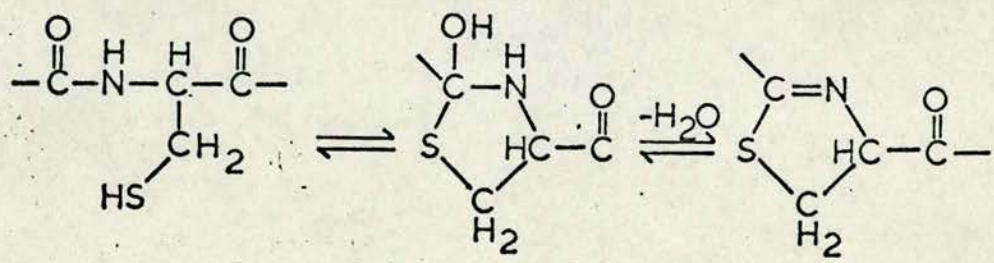
5 ml. aliquots of fresh stock solution were then treated with an equivalent quantity (to the TCA) of BuSH, shaken thoroughly and left to react for 4 hrs. After this time, they were dialysed against water as before. Once again, too, the external solution was concentrated and the spectrum obtained from it is shown in Fig. VI. 18. The peaks at 212nm and 240nm and the shoulders at 206nm and 230nm correspond exactly to peaks in the spectrum of $Au(CN)_2^-$. Unfortunately, the background absorption is high because of the phosphate and other impurities which were concentrated along with the gold complex.

Optical rotation measurements of the dialysates were made and these showed decreases of laevorotation of 11.2° and 24.6° in $[\alpha]_D^{18}$ and $[\alpha]_{436}^{18}$ respectively. The pH of the samples was measured also and found to be 7.00. From Fig. VI. 9a & b, it can be seen that these values correspond to the differences between native and reacted protein at pH 7.00, the extra disulphide bridge apparently having little or no effect on the specific rotation.

Whilst it was unfortunate that there was no more sensitive, readily available test for dicyanoaurite in solution, both sets of measurements pointed to the validity of the scheme proposed on page 113. It was therefore expected to find the TCA occupying a site close to the MMA one, from the crystallographic evidence.

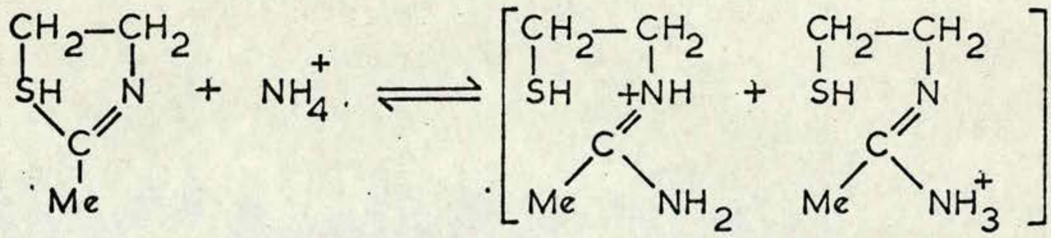
The Effect of Salt.

Since the overall object of this research was to produce heavy atom derivatives suitable for high resolution structure determination and since the crystals were grown by salting out, a study of the effect of salt on the binding of TCA to β-lactoglobulin was likely to be of use. Also, this was considered worthwhile in order to clarify the proposal put forward by Dunnhill (1963) that almost no reaction with pCMBS occurred in A.S. solution because of the formation of a thiazoline ring system involving the free sulphydryl group, the adjacent peptide carboxyl group and an NH_3 from some neighbouring group. His suggestion was founded in part on a measurement of the rates of reaction of β-lactoglobulin with pCMB as a function of pH (Dunnhill and Green (1965)). A plot of log k_2 versus pH showed a possible two-stage increase in the rate with increasing pH, there being a pause between pH 7.0 and 7.4. The second increase he proposed was the N - R conformational change but the first he put down to the breakdown of stability of a thiazoline ring. The ring formation on a peptide chain is shown below (see, for example, Calvin (1954). The mediation of the amine group was



derived from a paper by Calvin (1954) on the possible states of existence of glutathione in strong acid solution where the predominant form appears to involve a thiazoline ring.

Whilst the existence of thiazoline or other ring systems is not unknown (see, for example, the review by Cecil and McPhee (1959)), it was not seen how, of necessity, an $-NH_3^+$ group was involved. Its involvement was put forward as an explanation of the lack of reaction at high pH of pCMBS in the presence of 1.5M A.S. solution, because the ring's formation stabilised the low pH form. Linderstrøm-Lang and Jacobsen (1941) showed that 2-methyl thiazoline exists in solution as an equilibrium:



The addition of NH_4^+ to this system would force the equilibrium to the right thus destabilising the ring. If a parallel can be drawn between this system and that in β-lactoglobulin (and Dunnhill cites it in support of his theory) then the ring system is very unlikely in strong salt solution.

How, then, could the inhibition of the reaction with pCMBS be explained? Dunnhill showed that the Tanford transition was inhibited by strong A.S. solution as well as the reaction with mercury compound. Did the ammonium ion play a vital part? Work by Martin et al. (1959) on the hydrolysis of 2-methyl thiazoline showed that above pH 5.0 the open chain, N-acetyl-β-mercaptoethylamine was the more stable thermodynamically, the equilibrium

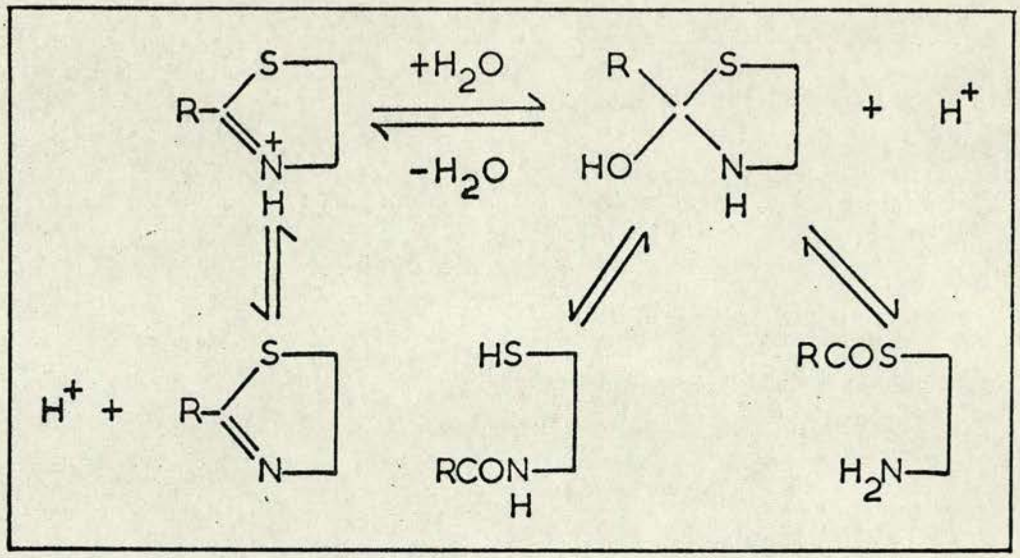


FIG. VI.10 Reaction scheme for the hydrolysis of 2-Methyl Thiazoline.

	RAW		CORRECTED	
pH	k_{obs}	Constant	k_{obs}	Constant
8.63	14.2	1860	14.2	1660
8.48	13.7	1816	13.7	1613
8.51	13.2	1834	13.3	1649
8.10	6.3	3124	6.3	2951
8.17	6.3	2833	6.2	2658
7.93	4.0	3135	3.9	2959
7.95	3.5	2990	3.5	2811
7.94	3.4	2806	3.4	2626
7.69	1.3	3585	1.3	3402
7.72	1.2	3424	1.2	3241
7.62	1.0	3286	0.95	3177
7.62	0.9	3237	0.84	3083
7.20	0.08	3928	0.082	3759
Mean intercept		3238 ± 327		3067 ± 330

TABLE VI.3 Uncorrected and corrected rate constants for the reaction of TCA with β-Lactoglobulin at various pH values.

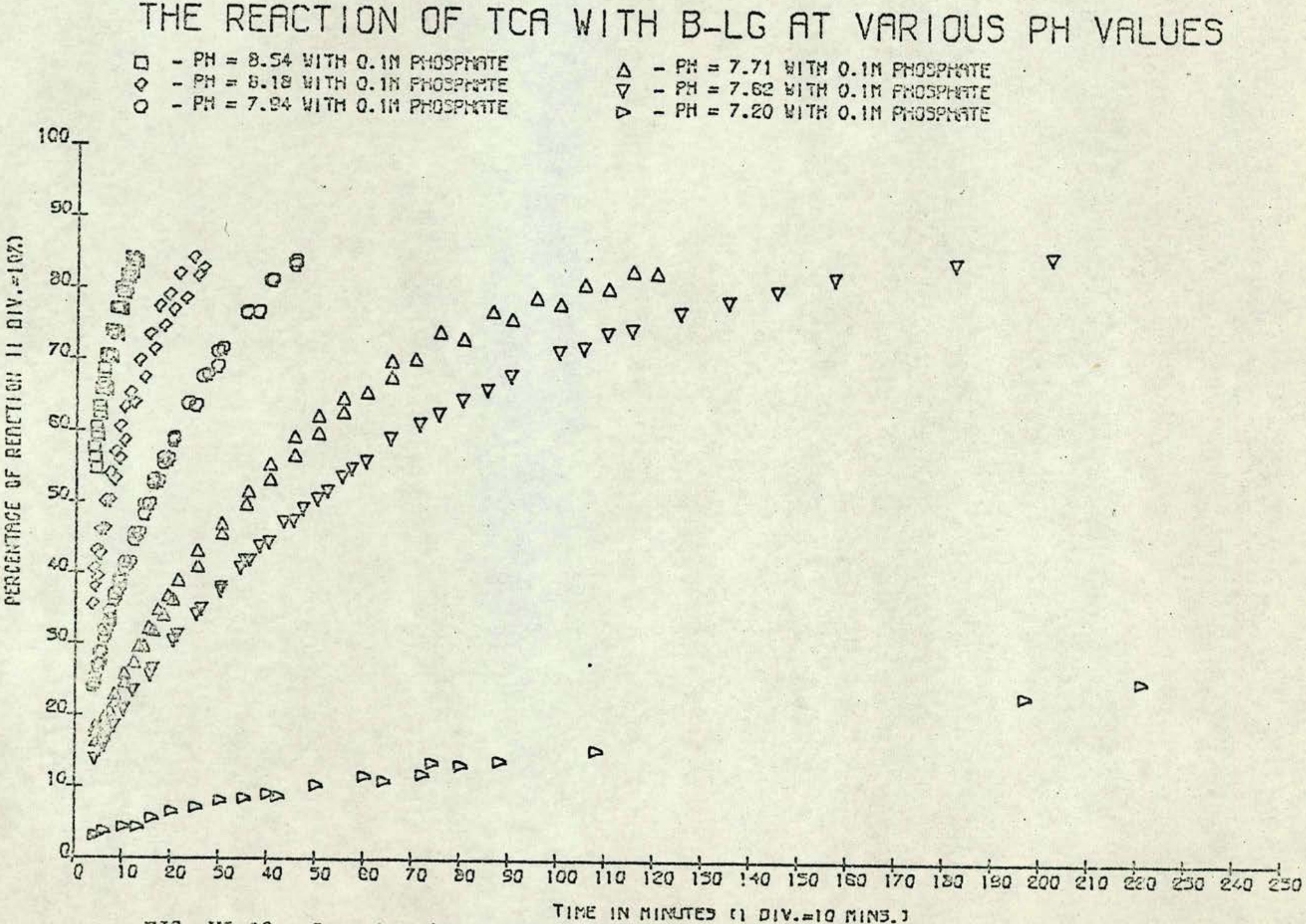


FIG. VI.20 Computer drawn graph of the percentage reaction versus time for TCA and β-Lactoglobulin at various pH values.

constant being 7×10^4 moles^{-1} l.$^{-1}$ at pH 7.0. Further, in alkaline conditions the ring, once formed, might be stable kinetically, and this is borne out by Schmir (1965) who examined the kinetics of hydrolysis of thiazolines between pH 8.5 and pH 12.7. He found a very slow, pH independent rate which indicated that the rate-limiting step in the hydrolysis was the formation of the thiazolinium ion, the necessary intermediate in the hydrolysis as shown in Fig. VI. 19. Any thiazoline ring system in β-lactoglobulin would appear to have to pass through an unstable region between pH 6.0 and pH 8.0, the region where the Tanford transition occurs. This can be seen to be the opposite of the proposal of Dunnhill.

For this reason as well as for the doubtful significance of the two-stage increase, the following kinetic study was undertaken. It was hoped to be able to show that the ammonium ions were not in themselves responsible for the decrease in reaction rate.

The Rate of Reaction of TCA with β-Lactoglobulin: Salt-Free.

The rate of reaction was measured by following the increase in absorption at 245nm. as a function of time. A 16:1 excess of TCA was used, this being added directly to the cuvette which already contained the protein solution at the appropriate pH and at 24.8°C. The measurements were made with the manual Unicam SP500 equipped with a thermostatted cell compartment connected to a waterbath the temperature of which was controlled to 24.8°C $\pm$ 0.1°C. Full experimental details are given in Appendix II. The values of k_{obs} are shown in Table VI. 3 and Fig. VI. 20 is a computer drawn graph of the percentage of reaction versus time.

The curves are very similar to those obtained by Dunnhill and Green (1965) save that the reactions are slower. Values were measured below pH 7.0 but they were too slow to be shown on the graph or even computed. At pH 6.57 complete reaction took about 240 hours whilst at pH 6.00 it was not even complete after 400 hours!

The intercept, c, was thought to be a measure of the solvation of the TCA which occurred very much faster than the reaction with the sulphydryl group. Trials with identical solutions but without the protein present, showed there was an optical density change of about 0.01 absorbance units associated with the introduction into the cuvette of the same quantity of TCA as was used in the rate determinations. Whether this value entirely accounted for the value of c was checked by recalculating the results after correction for the absorbancy change. The recalculated data showed almost identical values for k_{obs} and also for c (Table VI. 3). Because initially the speed of reaction at pH 8.51 is comparable with the solvation effects, the values of c obtained at this pH were considered to be unreliable. Neglecting these values, the remaining values of c were surprisingly constant and therefore some other effect involving both protein and TCA must be occurring. This could be a very fast non-covalent "association" of protein and ligand occurring at a specific site other than the sulphydryl.

The data in Table VI. 3 show that there is an increase in reaction rate with pH, similar to that found in β-lactoglobulin reacting with pCMBS, which corresponds to an increasing ease of

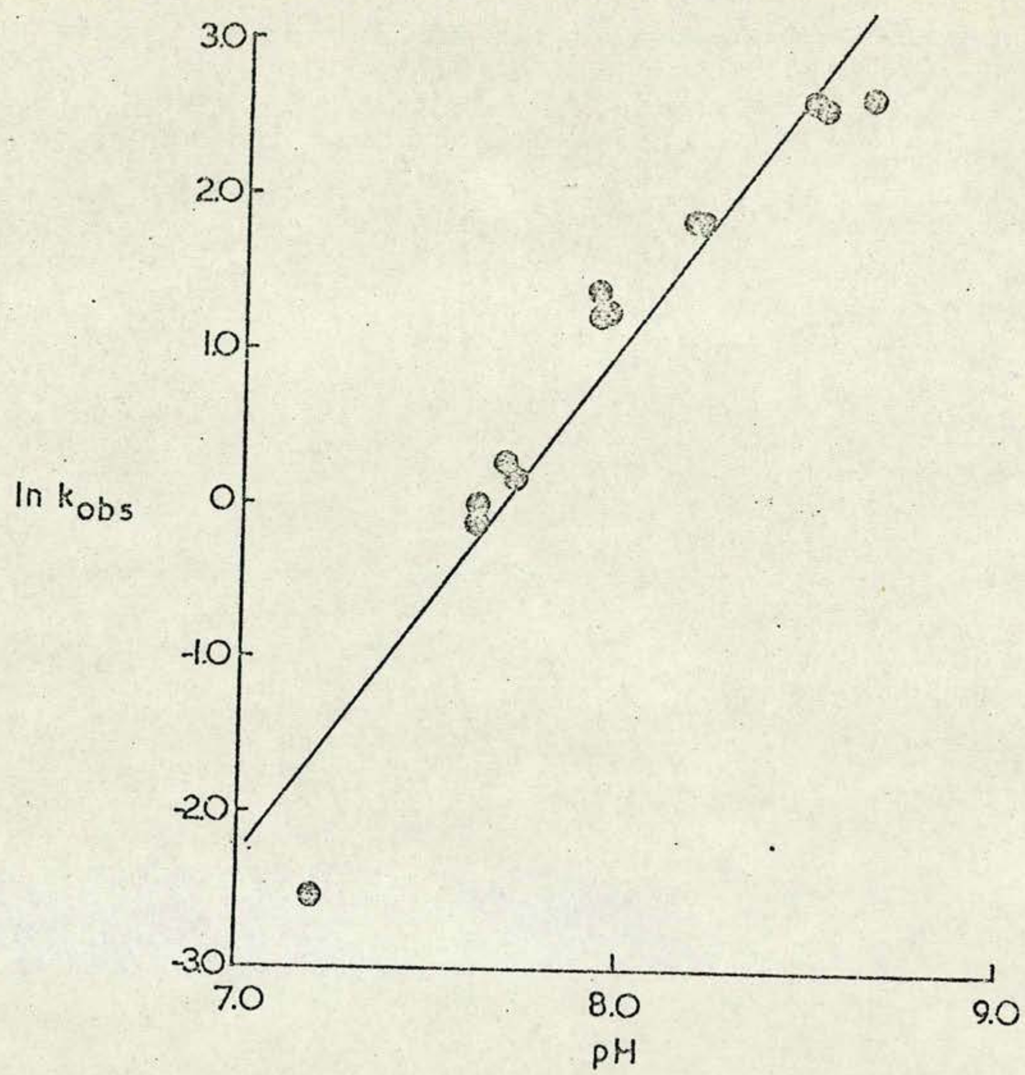


FIG. VI.21 Plot of log k_{obs} versus pH.

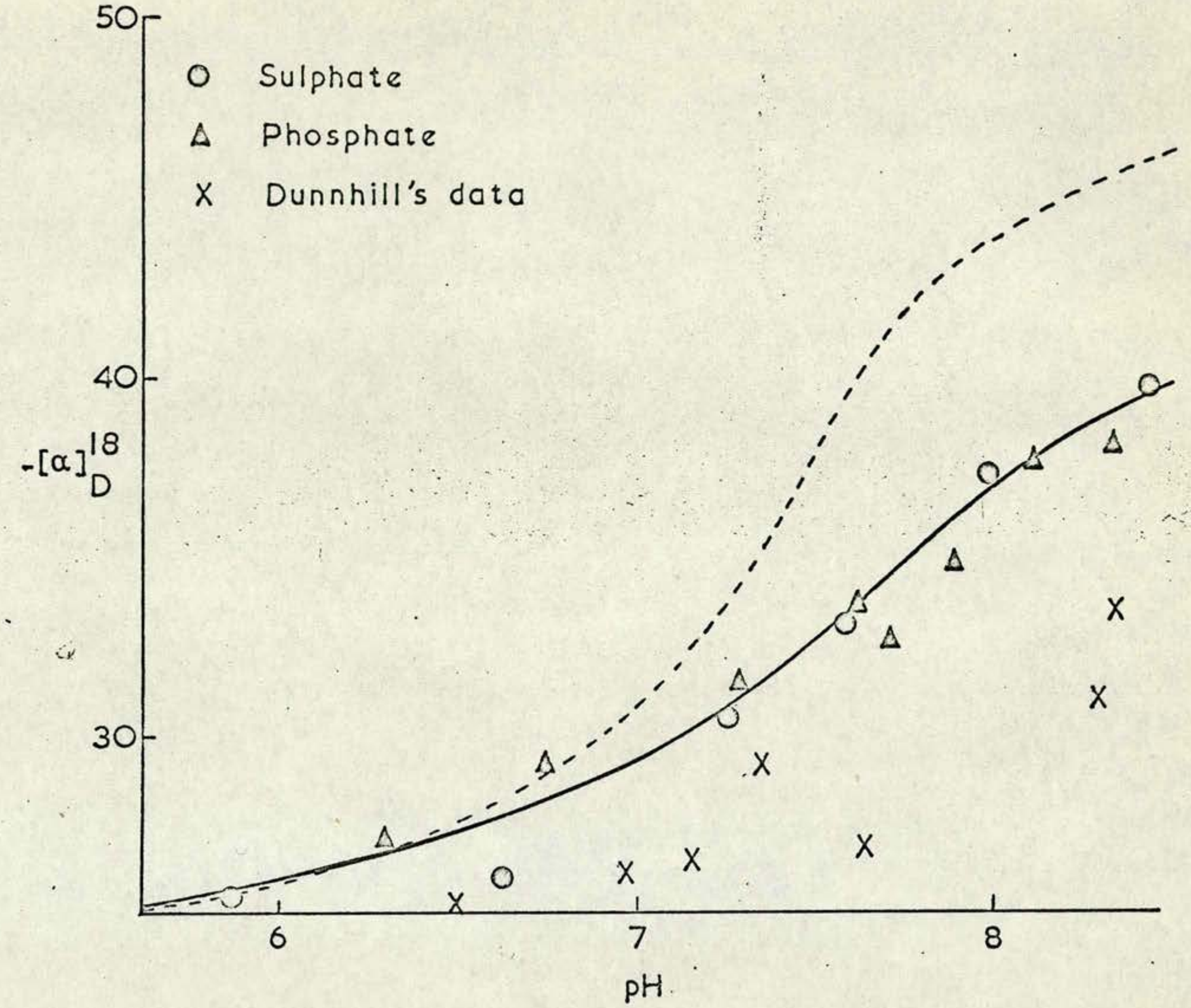


FIG. VI.22a Specific rotation at 589nm as a function of pH for 1.0×10^{-4} M protein in 1.5M sulphate and 1.0M phosphate. The data of Dunnhill and Tanford are shown for comparison.

reaction as the sulphydryl group becomes exposed by the conformational change. Over the range measured there was a steady increase in reaction rate so that a plot of log k_{obs} against pH gives a straight line (Fig. VI. 21) between pH's 7.2 and 8.5. This conflicts with the data of Dunnhill for the reaction of pCMBS with β-lactoglobulin. The pause in the increase of rate at about pH 7.4 obtained by Dunnhill was thought to be more likely due to experimental inaccuracy.

<u>The Effect of Salt on the Optical Rotation.</u>

The inhibition by A.S. of the Tanford transition was thought to be more likely caused by the binding of both NH_4^+ and $SO_4^=$ ions than by the extra stability gained from the formation of a thiazoline ring system, which, as has already been said, is more likely to be unstable at the pH's in question. Binding of ions to the surface of proteins is a well-known feature and is responsible for the salting-out effect. Semi-specific binding of ions to β-lactoglobulin would cause an electrostatic repulsion tending to oppose any change in conformation, leading to a diminution in the magnitude of the Tanford transition. If this were the case then any salt capable of salting out a protein should be capable of inhibiting the transition.

Because mixed phosphate buffer was known to produce crystals by salting out and because it contained no NH_4^+ (a necessary constituent for Dunnhill's proposal), it was the obvious choice to test the above postulation. Fig. 22 (a) shows the data for $[\alpha]_D^{18}$ plotted against pH for sulphate, phosphate and no salt (save 0.05M phosphate buffer, common to all the solutions) together

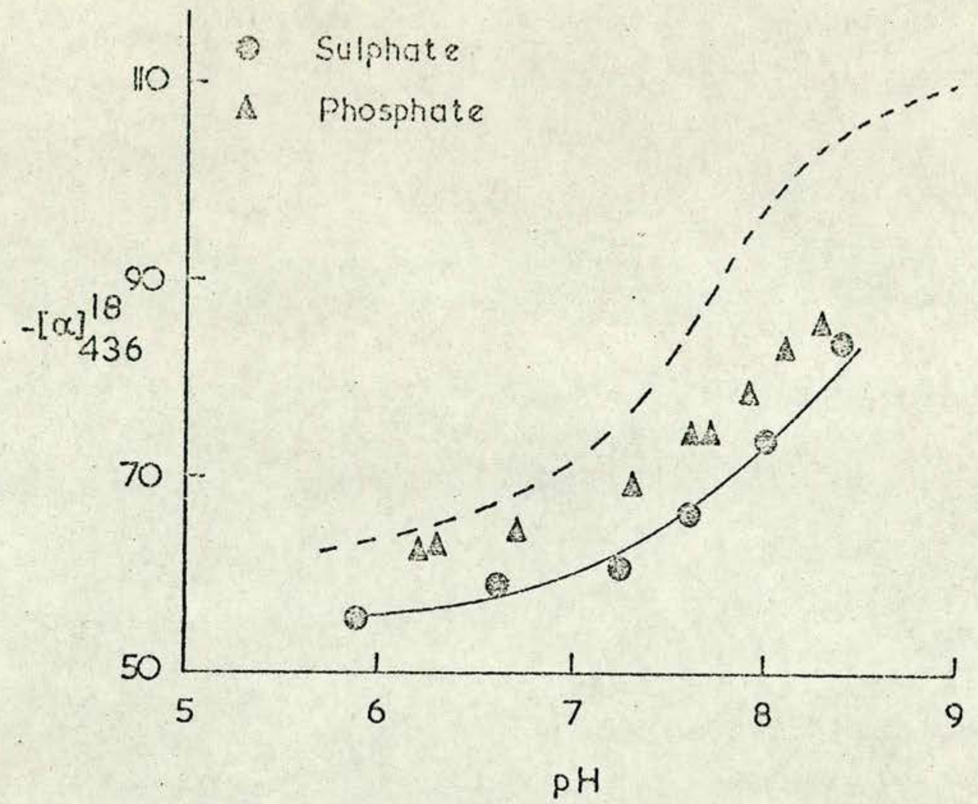


FIG. VI.22b As for Fig. VI.22a but at 436nm and the data for the native protein were determined in this work.

Salt	RAW k_{obs}	RAW Constant	CORRECTED k_{obs}	CORRECTED Constant
-	6.3	2070	6.3	2805
Sulphate	0.087	3902	0.093	3722
Sulphate	0.092	3873	0.100	3686
Phosphate	0.082	3764	0.086	3562
Phosphate	0.078	3600	0.078	3476
Chloride	0.077	3642	0.077	3421
Chloride	0.078	3600	0.075	3371
Chloride	0.078	3626	0.074	3403
Chloride	0.076	3681	0.073	3436
Mean intercept		3722 ± 106		3510 ± 124

TABLE VI.4 Uncorrected and corrected rate constants for the reaction of [illegible] with β-Lactoglobulin at pH 8.15 in various 1.0 M salt solutions.

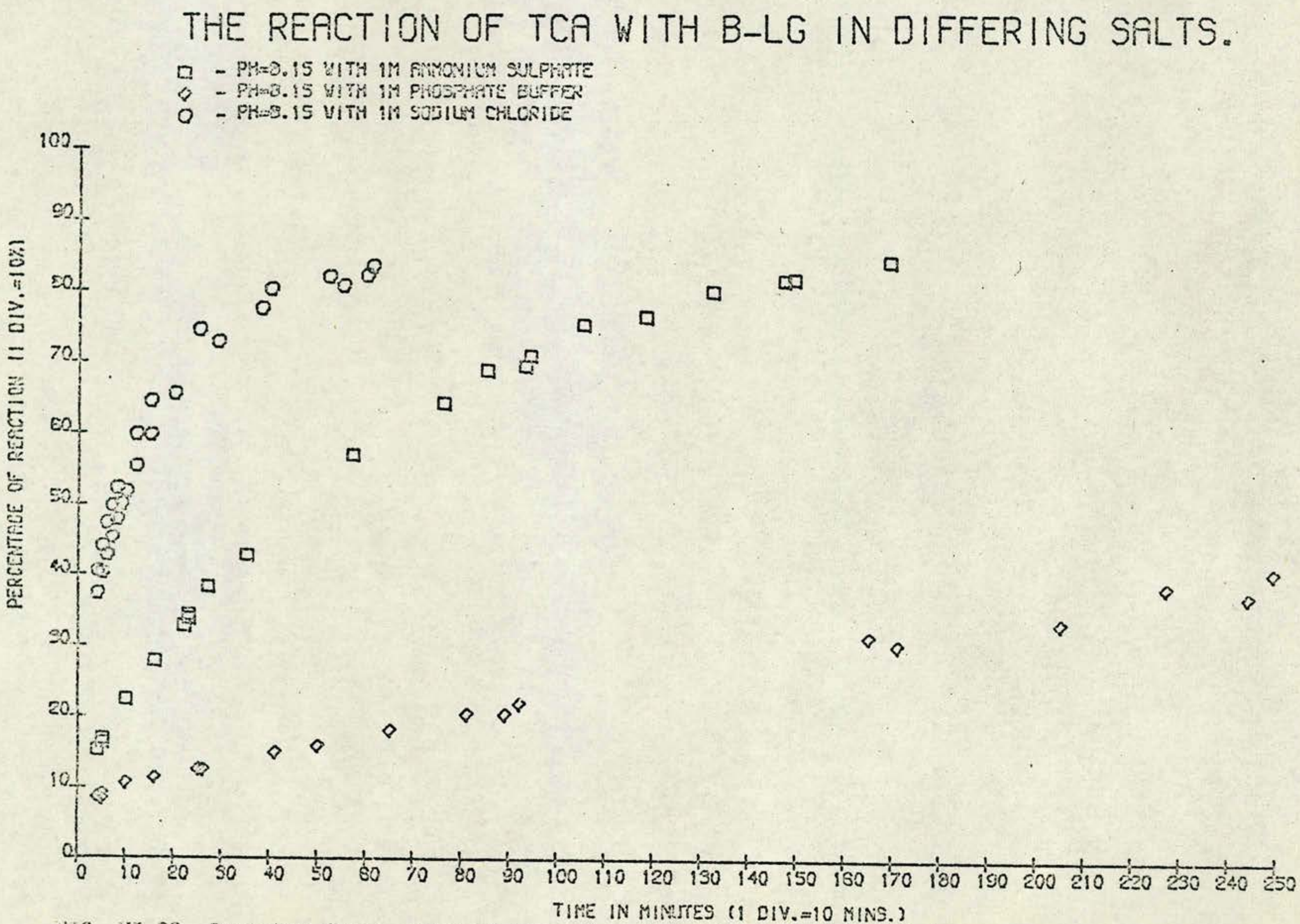


FIG. VI.23 Computer drawn graph of the rate of reaction of TCA with lactoglobulin in various molar salt solutions. The pH was 7.8.

with the idealised curve from Tanford. Dunnhill's data are also shown for comparison. A similar graph without the data of Dunnhill and Tanford is shown in Fig. VI. 22 (b) where $[\alpha]^{18}_{436}$ is plotted rather than $[\alpha]_D$ because the polarimeter gives a more intense beam and the optical rotation is greater at the lower wavelength thus allowing more accurate measurements. The phosphate and sulphate molarities were unfortunately different but the shape of the curves is identical showing that the effect of phosphate is at least similar to that of A.S., if not possibly greater (cf. data at 589nm.), in inhibiting the Tanford transition.

The Effect of Salt on the Rate of Reaction of TCA.

As another test of the lack of reaction being due to salt in general, rather than any particular ion, the rate of reaction in different salt solutions at the same pH was determined. The salts were A.S., phosphate and NaCl all at a concentration of 1M. The data are to be found in Appendix II and values of rate constant, together with that of the salt free run of equal pH are shown in Table VI. 4. Fig. VI. 23 shows the computer-produced plot of the percent reaction versus time for the three salts used.

Solvation in each of the three salt solutions was the same as far as could be detected on the SP500. The change in E_{245} was 0.02 absorbance units and this, as before, was used to correct for the solvation effect. These data are also shown in Table VI. 4. Once again there was a constant factor which did not appear to be merely solvation of the TCA by the salt

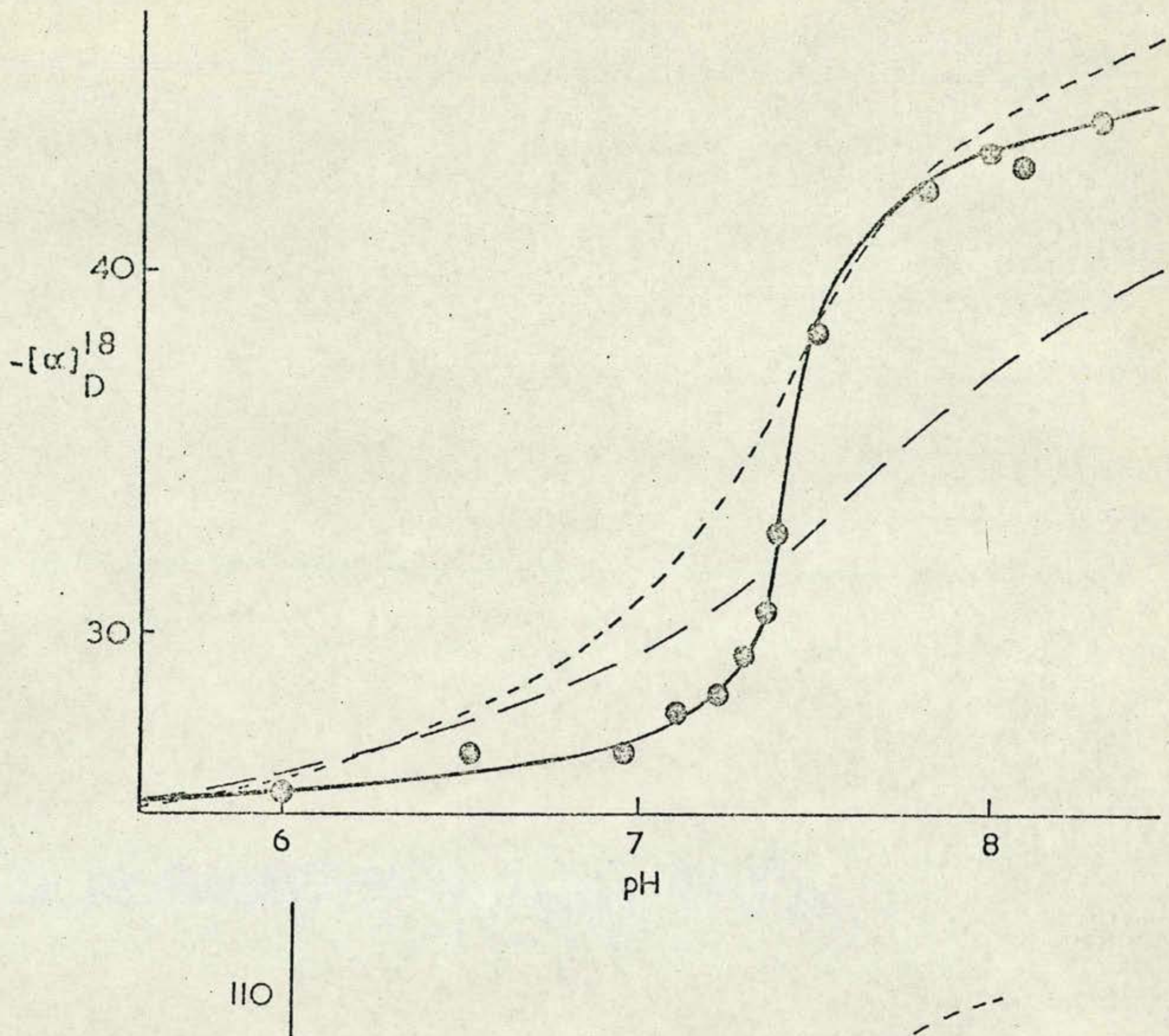


FIG.VI.24 The specific rotations at 539nm and 436nm for the protein-TCA complex as a function of salt in 1.5M salt. The protein was 1.001×10^{-4} M and there was a 5 : 1 excess of TCA.

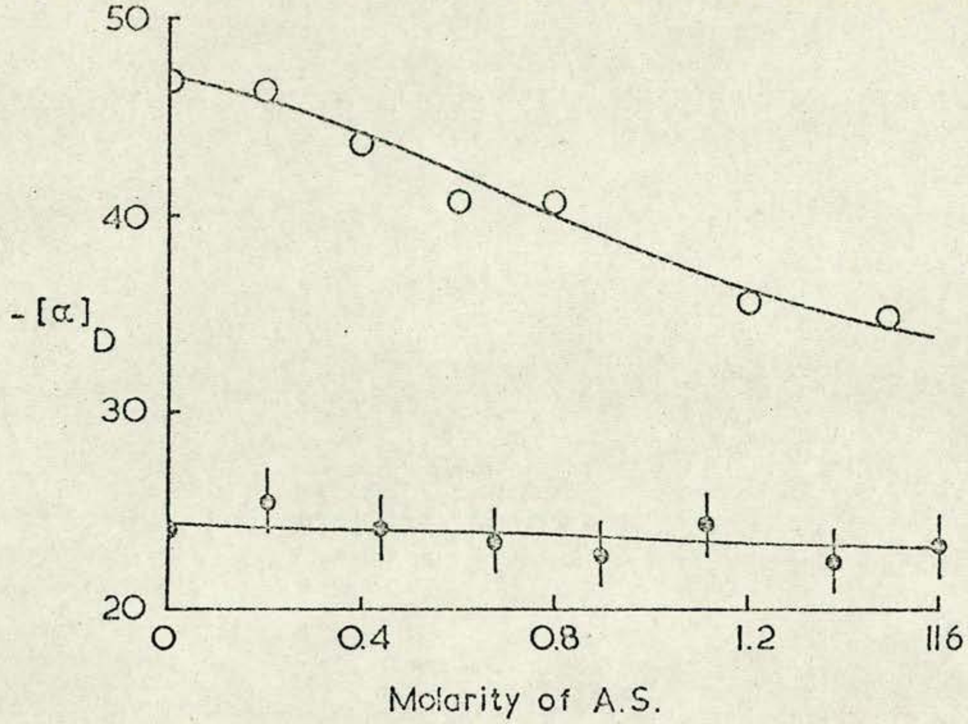
solution. In fact, it had been noticed that there was a quick rise of about 0.04 absorbancy units in the first 30 - 60 seconds after the addition of the TCA, followed by the slower, steady increase due to the TCA. This was especially noticeable in the measurements with salt where the rate was slower, being observed only when protein was present as well. Thus, it was concluded that the protein and TCA must be interacting in some way at a site other than the free sulphydryl one, because of this effect's being independent of pH.

It can be seen that the rate of reaction of TCA with the protein seems to be slowed down more by the increase of anionic charge rather than any specific action. This was the effect which was expected (Hofmeister (1887)) and it showed that the slowing down of the reaction with the free sulphydryl group in β-lactoglobulin in salt solution was not only dependent on the presence of NH_4^+ ions. Rather it depended on the presence of any salt which would associate with protein and/or ligand thus slowing the reaction down because of the need to remove the salt ions before reaction could occur.

<u>The Effect of Salt on the Optical Rotation of TCA-β-lactoglobulin.</u>

Salt was found to hinder the Tanford transition and to slow down the rate of reaction of TCA with the protein. What, then, was its effect on the optical rotation of the complex?

It might be expected that the specific rotation values for the complex would be lower than those in the salt-free state and, in fact, this was shown to be the case. The data are shown in Fig. VI. 24. It was noticed that the fully reacted protein in



<u>FIG. VI.25</u> Specific rotation of β-lactoglobulin as a function of A.S. concentration. The upper curve is of a 1.8mg/ml protein solution at pH 8.3. The lower curve is the data of Dunnhill for a 45mg/ml protein solution at pH 5.3. The errors are estimated.

salt solution had a conformation resembling much more closely that of the high pH side of the Tanford transition than did the conformation of the TCA-free protein. The salt-free data showed that a stable intermediate state between the R and S states could be readily formed by reaction with TCA. The data from concentrated salt solution now showed that this form corresponded quite closely to the R form in the salt free solution. One problem, of course, was that there could be a difference in the conformation in 1.5M salt to that in 2.0M salt from which crystals were grown. It was impossible, however, to increase the concentration above 1.5M in salt because the precipitation of the protein would have become noticeable. Indeed, filtration of the salt-protein solutions had to be carried out, even at 1.5M concentration, in order to obtain stable polarimeter readings. A plot of the optical rotation of the native protein versus the molarity of A.S. at pH 8.3 is shown in Fig. VI. 25. This is linear to 1.5M, but above this concentration does the trend continue or is there a levelling off, the conformation being locked in a stable intermediate state, corresponding to lattices Y and Z, between the N form corresponding to lattice X, and some other form given by reacting the protein with TCA and <u>then</u> crystallising it?

Green and his colleagues (unpublished work) have some evidence of crystallised β-lactoglobulin from a solution containing TCA in much the same way as for normal lattice Y crystals. If it crystallises in a different form from X, Y or Z, then this would appear to substantiate the above proposition that Y and Z are intermediate forms of β-lactoglobulin between the N and R states, which are stabilised by the presence of concentrated salt solution.

Conclusions from the Solution Work.

To summarise the findings of the last few sections, what has been found out from the solution work about the reaction of TCA with β-lactoglobulin?

First, it was shown that the TCA reacted with the free sulphydryl group of the protein as a replacement reaction,but this was shown to be a stable intermediate in the ultimate reduction of Au(III) to Au(I). The reaction was characterised by a change in the u/v spectrum of the protein at wavelengths less than 270nm. and also by an increase in the optical rotation to a point corresponding to a state intermediate between R and S.

Second, the proposal of the existence of a thiazoline ring formed and stabilised by the presence of NH_4^+ ions was shown to be incorrect for two reasons. First of all, whilst a thiazoline ring is reasonably stable at high pH, it has to pass through a region of low stability before attaining this state so that this cannot be the reason for the inhibition of the Tanford transition by A.S. Secondly, the same inhibition was shown to occur as a result of treatment with both mixed phosphate buffer and NaCl showing that the effect is not the result of NH_4^+ ions. However, the existence of a thiazoline ring as a tautomeric structure cannot be ruled out entirely, although the mediation of an ammonium ion rather than a sodium or potassium one in the formation of such a system can be.

Finally, optical rotation work coupled with X-ray work has shown that lattices Y and Z appear to be stable states between lattice X, the low pH form,and the conformation in the salt solution of the TCA-β-lactoglobulin complex at high pH. This latter form, if it exists, might well correspond to the upper state, R, of the Tanford transition. It is also interesting to notice that Mackenzie (1967) has found that β-lactoglobulin B in dilute buffer solution has a two-stage change in optical rotation when the pH is varied from 4.0 to 8.5, there being a pause between pH 6.0 and pH 7.0. This effect would tend to be moved to higher pH's with increasing salt concentration so that it might be a factor involved in determining the crystal form. However, pure β-lactoglobulin A does not have this plateau and it does form lattices X,Y and Z.

SECTION III

Three Dimensional Data Collection to 6Å Resolution.

As was mentioned already in section 1 of this chapter, the difference Fourier projections were unsatisfactory as a means of determining the whereabouts of the bound TCA molecules. In section 2 it was shown that one site should be the MMA one and it was hoped that a three-dimensional difference Fourier might be clearer and would bear out this prediction.

Strategy of Data Collection.

Lattice Y crystals have space group $B22_12$ thus having 222 diffraction symmetry. This means that two octants of the complete sphere of reflection will contain a complete diffraction spectrum, including the anomalous differences. Alternatively, one octant will contain enough information for a complete three-dimensional difference Fourier synthesis if the anomalous scattering is neglected. There are, in one octant, only 444 reflections to a resolution of 6Å.

Because the data were so few and in case the anomalous differences were needed to be used to determine the positions of the heavy atom sites, it was decided to collect a complete quadrant of the sphere. In the initial analysis both octants would be averaged giving a mean for the $|F_{PH}^{obs}|$ for which there were two measurements and also reducing the number of reflections requiring to be remeasured. (If anomalous scattering was going to be required, then the remeasurement would have to be extended to include both reflections.)

Data Collection.

Data were collected on a Hilger-Watts linear diffractometer using filtered Cu Kα radiation. The crystals were about $1 \times \frac{1}{2} \times \frac{1}{2}$ mm and the primary collimator was 0.8 mm in diameter. Full details of the theory and use of the linear diffractometer are given in the book by Arndt and Willis (1966) and the references therein. Briefly, the instrument is a device where a mechanical analogue is used to solve Bragg's Law by the Ewald construction. The crystal is mounted with each of two reciprocal axes, a* and b* say, parallel to one or other of the two slides, set at the angle between a* and b*. The real c-axis is then perpendicular to the slides and coincident with the goniometer axis. Then, the hk0 reflections can be measured automatically. An alteration of the l index requires tilting the frame of the instrument, at the same time moving the counter by equal amounts corresponding to $l.\lambda c^*$ for the equi-inclination setting. This layer can then be measured as before.

For $B22_12$ the axes are orthogonal which makes the reciprocal axes parallel to the real ones, thus making setting a crystal for measurement on the diffractometer relatively easy. Also, only reflections with h+l=2n are present so that the instrument's built-in skipping facility was used. For each reflection, the counter was stationary and the crystal was moved through a small angle, about $2\frac{1}{2}^{\circ}$, from one side of the reflection where a count of the background was taken, at constant speed through the reflection to the other side where, after output of the peak counts, another background count was made. After each step, the number of counts was printed out on a teletype, a punched paper tape also being produced. A one-minute oscillation motor was used so that the two background counts were of 15 seconds each and the scan through the peak, 30 seconds. At each reflection two complete cycles of counting were performed. The oscillation angle was found to be sufficient to allow for slight mis-setting of the crystal both on the goniometer and by the instrument and yet not so large as to reduce markedly the signal to noise ratio. The orthogonal axes meant that for single counter operation no special orientation was required for data collection on a linear diffractometer.

In all, three crystals were used to collect the data, the third one being used mainly for the collection of the low order terms which required setting by hand and for any others which were found either to have been measured incorrectly on both of the earlier crystals or not to have been measured at all. The data from this crystal were poor when compared with the photographic set. This could have been because of the crystal's shape causing more variation in the absorption with varying θ than the other crystals. It also had a greater fall-off. In this way, however, a complete data set was obtained.

As it happened, the a-axis of all three crystals was coincident with the goniometer axis. Thus, none of the h00 reflections could be measured by diffractometer since they were in a reflecting position no matter how the crystal was turned. These values had been obtained by photographic methods so that (at least) they (the data) had been measured. However, the fact that this row was coincident with the goniometer axis allowed a measure of the relative absorption to be made as a function of the angle the c-axis made with the X-ray beam. If the frame and counter are set in the equi-inclination position for an h00 reflection and

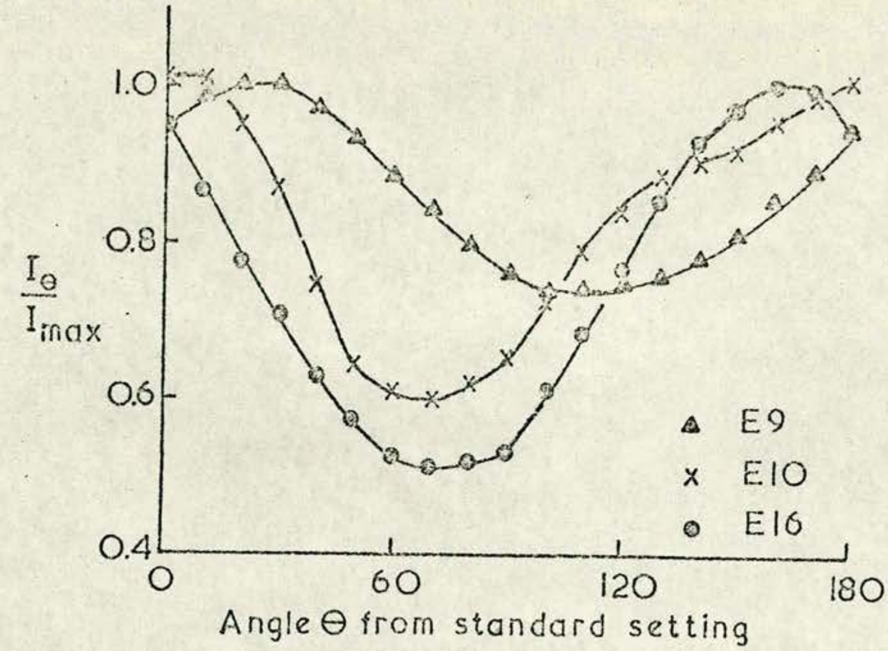


FIG. VI.26 The absorption curves for lattice Y, TCA-soaked crystals E9, E10 and E16.

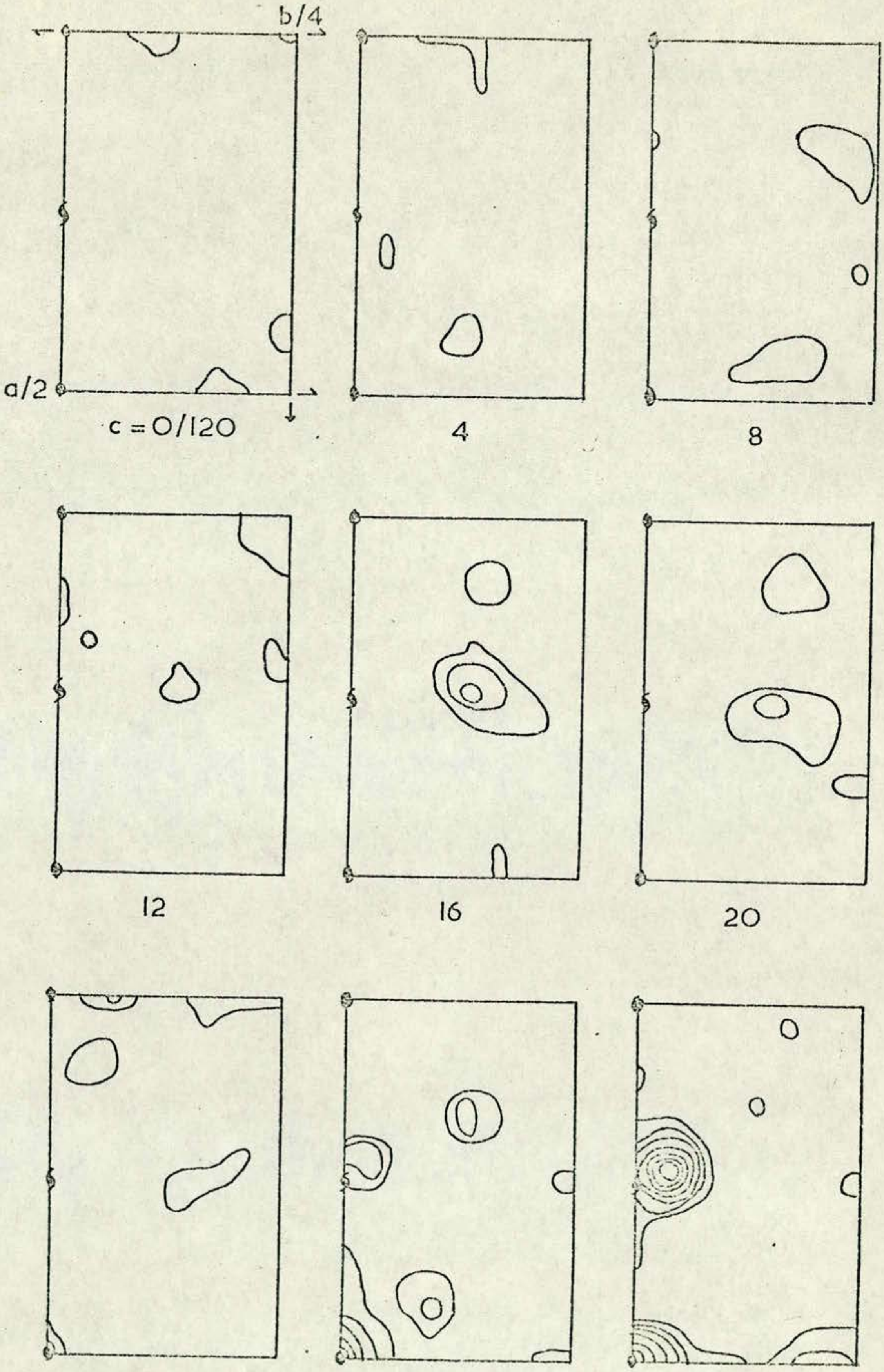
	Photo	1st Crystal	2nd Crystal	3rd Crystal
Value of agreement with photographs, D%	-	10.9	9.3	14.9
A	1.17	1.41	1.40	1.34
B	-8.15	-8.75	-8.38	-16.53

TABLE VI.5 Values of A and B for the three crystals used to collect the TCA derivative data together with those for the photographs. The final values of the agreement factor, D, for the crystals is also shown.

the frame rocked to and fro through the reflecting position by an auxiliary motor taking its timing from the oscillation motor, a measure of the h00 reflection can be obtained for the value of θ which one of the reciprocal axes makes with the X-ray beam. θ can then be altered by steps of 10^{o} or 20^{o} through a full revolution thereby obtaining a measure of the intensity as a function of θ. With the a-axis vertical, θ was defined in this case as the angle measured in a clockwise direction from the beam to the positive c^{*}-axis, initially parallel to the beam. The curve thus obtained can then be applied to the data as a semi-empirical absorption correction in the manner described by North, Phillips and Matthews (1968). In lattice Y, a conveniently strong reflection is the 400 and, accordingly, this one was measured in the manner just described. Fig. VI. 26 shows the scaled absorption curves for the three crystals. It would have been more accurate to obtain an absorption curve for each layer which was measured, but since the axial reflections only existed for every second layer and layer 9 was the highest to be collected, it was felt that the reflection 400 gave a fairly good approximation to the curves of the other layers.

Data Reduction.

The data were corrected for Lorentz, polarisation and absorption effects using a program written by Dr.D.W.Green based on the methods described for the treatment of linear diffractometric data by North (1964) and including checks on unequal background, mis-setting, negative counts and various teleprinter errors. COLLATOR was then used to compare each layer in turn


b/4
a/2
c = 0/120
4
8
12
16
20
24
28
32

with the native so that the fall-off constants could be obtained. These were applied to each layer, the layers were again compared with the native in order to check that the corrections were adequate and finally they were merged together to form a complete set. This set was then compared with the photographic set obtained earlier. The separate diffractometric layers were not scaled to the photographic set as a means of obtaining the scale factors because there were too few reflections in common to give accurate values for the scaling. The mean value of the agreement, D (see p.95), between the photographic and the counter data sets was 10.9% for the first crystal, 9.3% for the second and 14.9% for the third. Table VI. 5 summarises these values along with the scale factors. The complete data set obtained is listed in Appendix III together with the native data set (kindly made available by Dr. Green).

Three-Dimensional Difference Fourier Map.

As for the two-dimensional data, COLLATOR produced a list of $|F_H^{obs}|$ and α_P. These were used for calculating the three-dimensional difference Fourier synthesis, sections of which down the c-axis are shown in Fig. VI. 27. As can be seen, only two sites show up. Compared with the projections, the result was encouraging and indeed no account needed to be taken of the anomalous dispersion in order to obtain a clear picture of the sites of substitution. The fractional coordinates for the major peak, A ,and minor peak,B were measured as

A: 0.250 , 0.019 , 0.267

and

B: 0.500 , 0.000 , 0.270

Peak A was seen to be close to the MMA site thus apparently bearing out the proposal at the end of the previous section. Peak B was found to be close to the HGI site, the one it was hoped might be the major site.

Because of the possibility of some carry-over from the sites used to calculate the phases, it was thought advisable to produce a $(\Delta F)^2$ synthesis as a check. This was done in three dimensions and it showed a prominent peak corresponding to the A-B vectors at about $(\frac{1}{4},0,0)$. The B-B' vectors were close to the origin and were masked to some extent by the diffraction ripple of the origin peak. The A-A' vectors appeared at about $(\frac{1}{2},0,0)$. All of these were consistent with the two peaks, A and B, obtained in the Fourier map and no other peaks were found in the $(\Delta F)^2$ map. This showed that the peaks were in fact a result of the TCA, the larger being the predicted one at or around the MMA site. The next stage was to refine the sites' parameters in order to obtain some idea of the occupancy of each site and the best positional parameters for them.

Refinement of the TCA Parameters.

Because $B22_12$ has three mutually perpendicular centrosymmetric projections, it was possible to refine all the positional parameters using only the centric data. The refinement was carried out using Hart's method by means of the program, PANGLOSS, described in Appendix II. The scattering factor curve used was that of gold which was considered a fair approximation at 6Å resolution since the carbon and nitrogen atoms were spread out and would probably be indistinguishable from the bulk of the protein. The occupancy was set arbitrarily

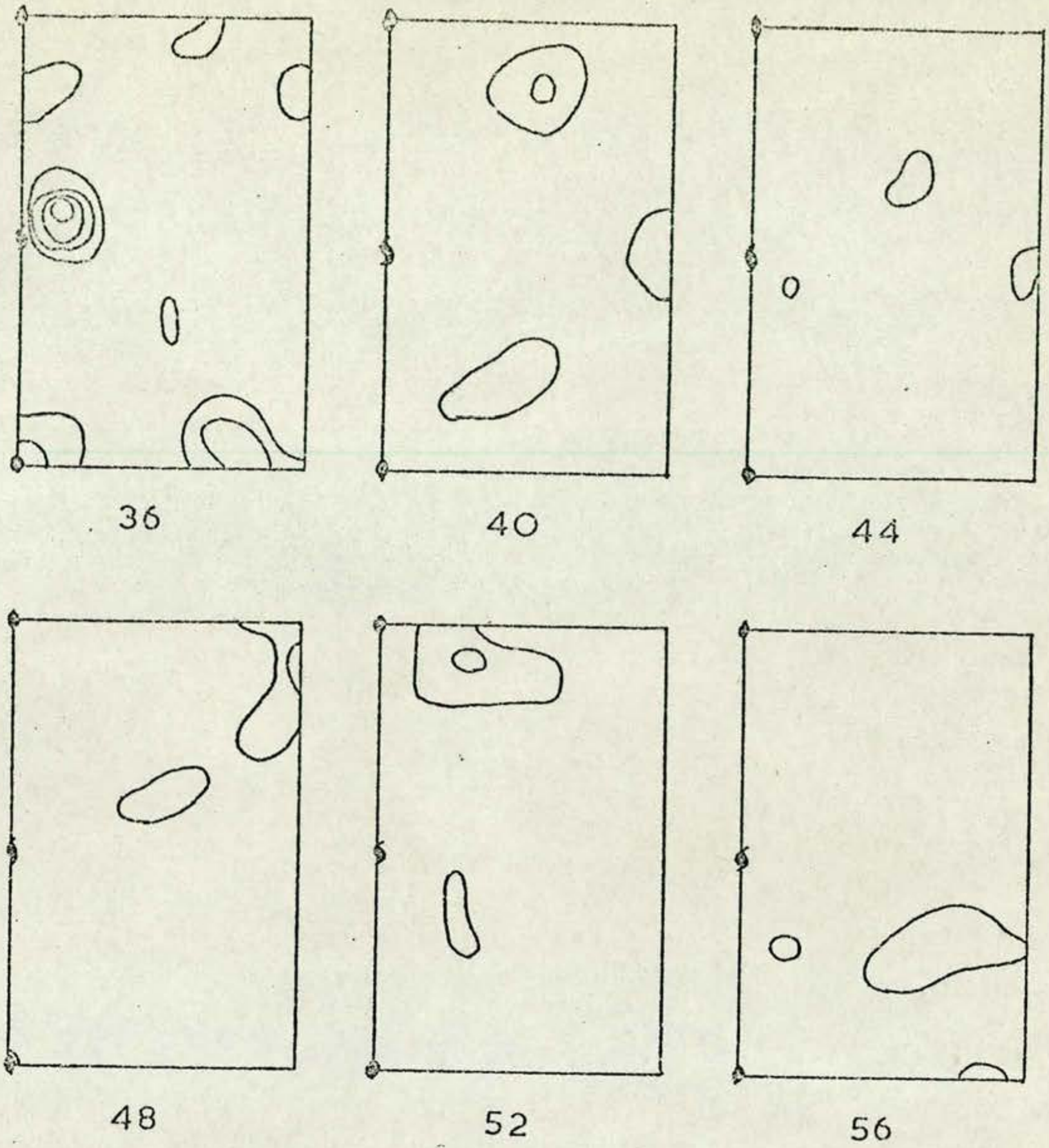


FIG. VI.27 Three dimensional difference Fourier of the TCA derivative with sections perpendicular to the c-axis. The negative and zero contours have been omitted and the interval between the positive contours is 0.05 electrons/$\mathring{A}^3$.

with the native so that the fall-off constants could be obtained. These were applied to each layer, the layers were again compared with the native in order to check that the corrections were adequate and finally they were merged together to form a complete set. This set was then compared with the photographic set obtained earlier. The separate diffractometric layers were not scaled to the photographic set as a means of obtaining the scale factors because there were too few reflections in common to give accurate values for the scaling. The mean value of the agreement, D (see p.95), between the photographic and the counter data sets was 10.9% for the first crystal, 9.3% for the second and 14.9% for the third. Table VI. 5 summarises these values along with the scale factors. The complete data set obtained is listed in Appendix III together with the native data set (kindly made available by Dr. Green).

Three-Dimensional Difference Fourier Map.

As for the two-dimensional data, COLLATOR produced a list of $|F_H^{obs}|$ and α_P. These were used for calculating the three-dimensional difference Fourier synthesis, sections of which down the c-axis are shown in Fig. VI. 27. As can be seen, only two sites show up. Compared with the projections, the result was encouraging and indeed no account needed to be taken of the anomalous dispersion in order to obtain a clear picture of the sites of substitution. The fractional coordinates for the major peak, A ,and minor peak,B were measured as

A: 0.250 , 0.019 , 0.267

and

B: 0.500 , 0.000 , 0.270

Peak A was seen to be close to the MMA site thus apparently bearing out the proposal at the end of the previous section. Peak B was found to be close to the HGI site, the one it was hoped might be the major site.

Because of the possibility of some carry-over from the sites used to calculate the phases, it was thought advisable to produce a $(\Delta F)^2$ synthesis as a check. This was done in three dimensions and it showed a prominent peak corresponding to the A-B vectors at about $(\frac{1}{4},0,0)$. The B-B' vectors were close to the origin and were masked to some extent by the diffraction ripple of the origin peak. The A-A' vectors appeared at about $(\frac{1}{2},0,0)$. All of these were consistent with the two peaks, A and B, obtained in the Fourier map and no other peaks were found in the $(\Delta F)^2$ map. This showed that the peaks were in fact a result of the TCA, the larger being the predicted one at or around the MMA site. The next stage was to refine the sites' parameters in order to obtain some idea of the occupancy of each site and the best positional parameters for them.

Refinement of the TCA Parameters.

Because $B22_12$ has three mutually perpendicular centrosymmetric projections, it was possible to refine all the positional parameters using only the centric data. The refinement was carried out using Hart's method by means of the program, PANGLOSS, described in Appendix II. The scattering factor curve used was that of gold which was considered a fair approximation at 6Å resolution since the carbon and nitrogen atoms were spread out and would probably be indistinguishable from the bulk of the protein. The occupancy was set arbitrarily

	Initial	After 22 cycles	After 26 cycles
Occupancy (electrons)	98	91	-
B	20	-	20.008
x	0.25	0.25882	-
y	0.19	0.03803	-
z	0.267	0.2610	-
Occupancy (electrons)	79	60	-
B	20	-	20.03
x	0.50	0.50101	-
y	0	0.03162	-
z	0.27	0.27825	-
Scale factor	1.000	0.976	-
R.M.S. Deviation	325	184	184
R	0.550	0.449	0.449

Radial Distribution

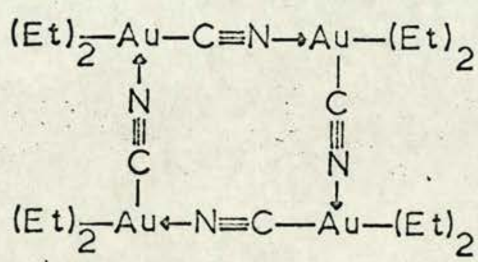
Zone	Number	$\overline{d^*}$	$\overline{F}_p$	$\overline{F}_{ph}$	$\overline{F}_h^{obs}$	$\overline{F}_h^{calc}$	$\overline{Del}$	$\overline{R}$
0	20	.0406	582	623	370	365	148	.406
1	23	.0814	604	645	400	379	174	.471
2	22	.1049	779	657	308	267	140	.522
3	25	.1237	619	668	285	225	170	.757
4	22	.1417	422	403	207	266	113	.426
5	23	.1548	365	411	284	272	104	.383
6	9	.1646	419	446	207	244	139	.568

TABLE VI.6 Refinement stages and final radial distribution of the two TCA sites in lattice Y.

at 98 electrons for the major site and at 70 electrons for the other one, these values being the relative peak heights from the Fourier map with approximately full occupancy at the MMA site. The isotropic temperature factor was not refined but left at a value of 20 for both sites throughout. The scale factor was refined separately after the first few cycles of refinement of the other parameters and then all parameters were allowed to refine for the final few cycles. The values of the parameters at the various stages of refinement are shown in Table VI. 6 along with the final radial distribution of structure factors.

The value of R, the agreement between observed and calculated values of F_H can be seen to be rather high when compared with typical, "good" values of between 0.25 and 0.35 (Cullis et al. (1962)). This indicated that the sites used to calculate the heavy atom contribution were not sufficient to account for the observed changes but, since there was known to be a fair degree of non-isomorphism, this was to be expected.

The final occupancies of the two sites corresponded to 114% of a gold atom at the free sulphydryl group and 75% of a gold atom at the other site. Comparison of the refined sites of TCA with those of MMA and HGI showed the following. The distance between the Hg atom of MMA (which may be assumed to be what shows up at 6Å resolution) and the Au atom of TCA is about 1.3Å, whilst that between the Au and 'Hg' atoms at the other site is some 4.5Å. Clearly some more specific sort of binding in a possibly more rigid part of the molecule is occurring at the former site compared with the latter. This is in accordance with the solution studies which showed that there was binding at the free sulphydryl group. The minor site in the TCA derivative is situated between the subunits and here the difference in shape between the TCA and $HgI_4^{=}$ ions would probably be sufficient to account for the fairly large shift. The TCA, being planar, could slip further in between the subunits as is indicated by the closeness of the gold to the dyad. The Au-Au distance across the dyad is about 4.4Å which is slightly less than half that between the HGI sites. On page 90 it was suggested that there might be a bridged-type of compound but the only one discovered has been the square one shown below (Phillips and Powell (1939)). The Au-Au distance is about 5.2Å and four gold atoms are involved



Et : Ethyl

The planes of the two TCA molecules must therefore be twisted out of the plane perpendicular to the dyad since the nearest approach of the gold atoms assuming a van der Waal's radius for the nitrogen atom of 2.2Å is about 8.5Å. This point can only be properly cleared up when the structure is fully known.

The refinement program also assigns calculated phases (signs) to the observed intensities so that a difference Fourier projection,

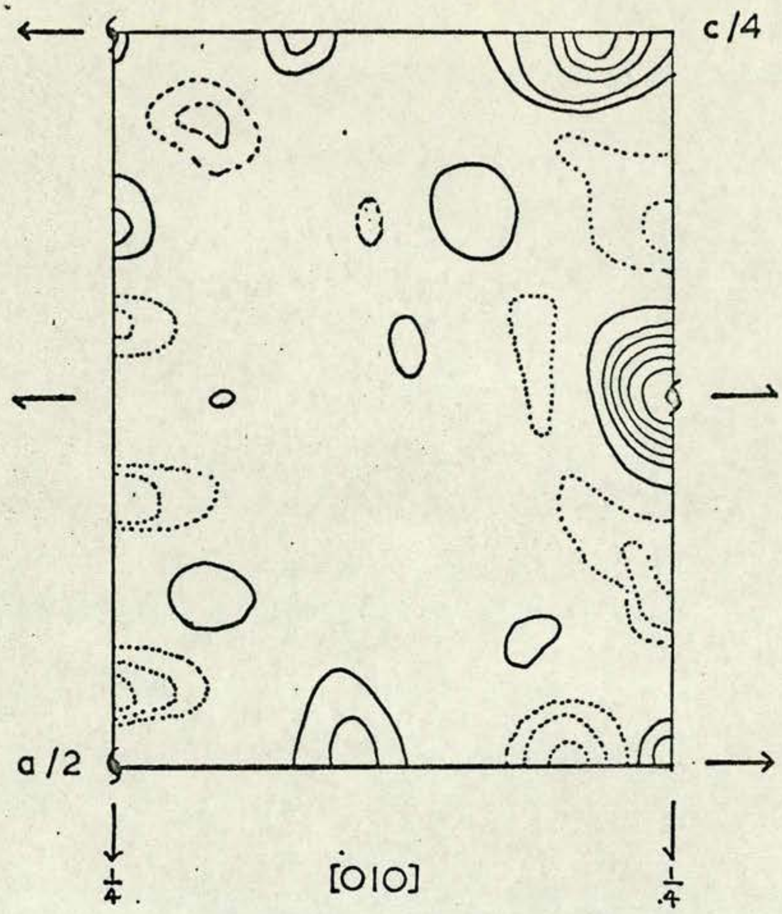


FIG. VI.28 The [010] difference Fourier projection of the TCA derivative using the signs calculated by PANGLOSS. The contours are at 2 electrons/$Å^2$ intervals.

	TCA	HGI
Site 1		
Occupancy (electrons)	62	232
x	0.066	0.102
y	0.405	0.448
z	-0.007	-0.002
Site 2		
Occupancy (electrons)	112	137
x	0.016	0.016
y	-0.168	-0.162
z	-0.146	-0.056

TABLE VI.7 Comparison of the TCA and HGI sites in lattice Z crystals.

independent of the native phases, can be produced. Such a Fourier map is shown in Fig. VI. 28 where it can be seen that the spurious peaks have disappeared leaving only the two heavy atom sites. Further, a 'double-difference' projection showed there to be no other significant features. A comparison of the signs used in these maps and the similar ones for the other two projections with those of the native showed that some 40% of the signs had changed. This figure includes the low order terms affected by salt so that if more, higher resolution data were available this fraction of sign changes would most probably decrease rapidly. However, this explains the uninterpretability of the projection data as being a result of approximately half of the signs being incorrect.

Lattice Z and TCA.

Paralleling this work, Green has examined lattice Z crystals soaked in TCA. A reappraisal of the h0l projection data collected from these crystals showed two things. First, as for lattice Y, the cell dimensions had changed by a moderate amount.

Native: $a = 54.4 \pm 0.4Å$; $c = 113.2 \pm 0.6Å$

Derivative: $a = 54.2 \pm 0.5Å$; $c = 116.1 \pm 0.8Å$

Second, and more surprisingly, the sites of substitution were found to be close to the major and minor sites of HGI with no substitution at the MMA site. Furthermore, the minor HGI site was now the major TCA site. Table VI. 7 summarises the positions and occupancies of both HGI and TCA derivatives of lattice Z. But why was there no reaction at the MMA site as predicted by both solution and X-ray work on the Y crystal form? MMA is known to attach to the sulphydryl site in lattice Z so that the thiol group cannot be completely shielded by the protein. The MMA molecules, however, is long and thin and much more "thiophilic" than the cross-shaped TCA. This, together with the different packing arrangment which seems to increase the availability of the HGI minor site in Z, could be the explanation of the lack of reaction at the sulphydryl site.

Because of the lack of isomorphism of the lattice Z TCA derivative, it could not be used for phase determination as it had been prepared for the above work. The ratio of heavy atom to protein had been about 3:1 but, if a fresh preparation using perhaps a 1:1 ratio were set up and allowed a longer diffusion time, the result might be a derivative with only the HGI minor site occupied.

Conclusions from the TCA Work.

To sum up, what can be learnt from the work with TCA and β-lactoglobulin? As far as lattice Y is concerned, it is encouraging that the solution and X-ray work tie up so nicely. But from the point of view of obtaining a better derivative than that using $HgI_4^{=}$ for high resolution studies, no such success was achieved. This bears out the earlier comments on the difficulty of preparation of suitable derivatives because of the uncertainty about which peptide side chains are available for binding to any added ligand, the lack of knowledge about the effects of solvation on both protein and ligand,and whether the ligand is actually bound chemically to,or merely associated with a particular site.

As was stated above, a fast "solvation" involving the protein and TCA occurs to approximately the same extent, as far as can be judged from the spectrophotometric evidence, regardless of the aqueous solvent. Earlier, too, the possibility of this being the result of the TCA's reacting at the HGI site was mentioned. In solution, however, packing arrangements need not be considered to the same extent as they must in the crystalline state so that the reaction could equally well be with the HGI minor site, when it is remembered that it is this site which has the greater occupancy in lattice Z.

The change in cell dimensions on the inclusion of TCA is interesting. The HGI sites in lattice Y in the same dimer are some 9Å apart, across the molecular dyad. Because this dyad is also the crystallographic one along c, that suffering the greatest change, a bridging compound pulling the subunits together cannot be the reason for the shortening of cell dimension. In lattice Z, the expansion is along the c-axis which is not the molecular dyad. In other words, the different sites of substitution appear to be responsible for differing effects on the cell parameters. Now, in lattice Y, the sulphydryl group is known to be the major site of substitution and, from solution, reaction at this site is known to result in a specific change in conformation. It is therefore possible that the contraction in cell dimension in Y is a direct consequence of the protein's attempt to change to another conformation hampered by the constraint of the crystal lattice. This other form could give a different form when crystals are grown from a protein-TCA mixture and would be worth investigating.

One question which has not been dealt with is the possibility of specific complexes with groups other than thiols. $PtCl_4^{=}$ has recently been found to bind to methionine residues where these are exposed and even to associate with histidines and disulphide bridges (Dickerson et al. (1969)). Another equally important question is that of the substitution of the complex's own ligands by the solvent or salt ions present in solution. Sigler and Blow (1965) discuss this point with reference to $PtCl_4^{=}$ and α-chymotrypsin. In other words, the heavy atom which is bound to the protein is not that which is added. Such changes require the complete structure to be elucidated before a full answer can be obtained.

As a final point, it would be of interest to examine the effect of TCA as a sulphydryl reagent on other proteins. Whilst mercurials are the recognised reagents for such investigations, from the X-ray point of view TCA would appear to be a better anomalous scatterer thus affording easier determination of the enantiomorph. Another useful point is that it is a different shape from most of the recognised mercurials, the four coordinate ones being tetrahedral, so that it can be used as another probe, in solution, as to the availability of the thiol group and the size of the hole in which it is situated.

CHAPTER VII

β-Lactoglobulin and Tetracyanoplatinite (II)

Having found that TCA was not going to be of use as a replacement for $HgI_4^=$ in lattice Y, it was decided to examine the binding of tetracyanoplatinite (II) (TCP). This, like TCA, is a d^8, square-planar complex but unlike the gold compound, it is already in its lower (non-zero) oxidation state. As can be seen from the oxidation potential calculated from the stability constant, β_4=41 (Chemical Society Special Publication No. 17),the tetra-substituted ligand should be moderately stable under normal conditions.

$$2e^- + Pt(CN)_4^= \rightleftharpoons Pt + 4CN^- \qquad E = -0.01v$$

This value is subject to fairly large errors, however, since the value of β_4 could be in error by as much as 10% and the value of the potential for Pt/Pt^{2+} (Latimer (1952)) is only approximate. Nevertheless, because of this, it was thought possible that the likelihood of reaction at the MMA site would be diminished and that the preferred binding site might be the desired HGI one.

The size of the TCP molecule is almost identical to that of TCA. Wykoff (1965) quotes a rather abstruse paper on the structure of $KNaPt(CN)_4.3H_2O$ which gives the values for the Pt-C and C-N bonds as 1.93Å and 1.16Å respectively. These are almost the same as the equivalent bond lengths in TCA. Apart from the similarity in size, there was another factor the effect of which could not be estimated so easily. This was the differing charge on the complex. Like $HgI_4^=$, the TCP has an overall charge of -2 and could this also be a factor in forcing the complex away from the MMA site?

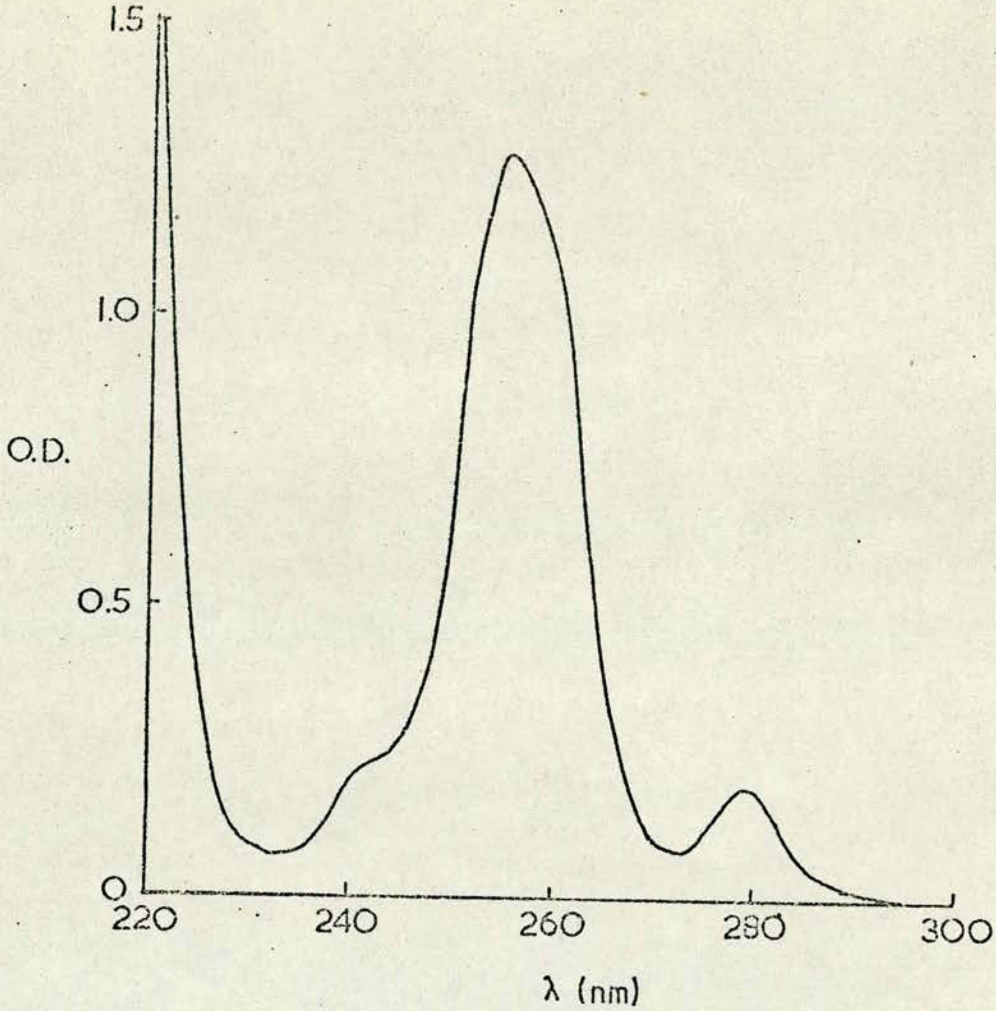


FIG. VII.1 Spectrum of a 4.3×10^{-5} M solution of $(NH_4)_2Pt(CN)_4$ at pH 7.6.

	Diffractometer Data	Photograph/Diffractometer Data	TCA Data
A	1.303	1.255	1.41
B	-13.22	-11.34	-8.62
D, %	8.0	10.2	10.1

TABLE VII.1 The fall-off correction terms for TCP showing the effect of the photographic data set. The data for TCA are included for comparison.

The first step was to examine the solution properties as before, using spectrophotometry and polarimetry to see if any reaction could be detected.

Solution Studies on TCP with β-Lactoglobulin.

First of all, the u/v spectrum was measured. This is shown in Fig. VII. 1 and the maxima agree well with those reported by Parumareddi, Liehr and Adamson (1963). As can be seen, there is an intense absorption band between 250 and 260nm with a smaller one at 280nm. These unfortunately coincide with the protein bands in such a way as to make accurate measurement of the spectra, either direct or difference, almost impossible in this region. Thus, the only way of telling if any reaction had occurred was from the optical rotation.

A stock 0.026M solution of the ammonium salt of TCP was made up in 0.1M mixed phosphate buffer, pH 7.8. Almost all of the solid dissolved but a small amount remained which was filtered out. This may well have been $K_2Pt(CN)_4$ which is slightly less soluble than either the sodium or ammonium salts (C.R.C. Handbook 47th Edition). The concentration was determined using the value for the molar extinction coefficient given by Parumareddi et al. (1963) as 1590 cm^2/mole. It was found to be 0.021M and it was this solution which was used in the subsequent experiments.

Solutions were made up in 10ml standard flasks as shown below.

5.00mls protein solution, 2.01×10^{-4}M in 0.1M phosphate

0.20mls ligand, 0.021M, omitted in the reference

4.80mls water 5.00 mls in the reference

This gave a molar ratio of 4:1 TCP to protein and the flasks were set aside at 3°C for 24 hrs. to equilibrate. The pH of each solution was measured after the optical rotation measurements had been made and was found to be 7.80 in all cases. The mean specific rotations at both 436nm and 589nm are shown below together with that of the reference.

	$-[\alpha]_D^{18}$	$-[\alpha]_{436}^{18}$
Protein alone	44.4°	96.2°
Protein-TCP	44.4°	97.2°
Error	±0.1°	±0.2°

There was no effect, therefore, similar to that in TCA. This was confirmed when the solutions were measured a week later after storage at 3°C and values were obtained which were not significantly different from those above.

Thus, it appeared that no binding of the type described for TCA was taking place or, in other words, the TCP was not binding to the free sulphydryl group, as far as could be seen from the optical rotation results. This was encouraging and the next stage was to examine lattice Y crystals which had been treated with TCP.

X-Ray Work on Lattice Y Crystals Soaked in TCP.

Crystals of lattice Y with a heavy atom to protein ratio of 2:1 had been found to give no changes and those with a ratio of 6:1 had been found to be non-isomorphous although large changes had occurred (Green, unpublished observations). A preparation with a ratio of 4:1 had been prepared in London but had never been examined. Therefore, before preparing any fresh crystals these ones were checked for isomorphism and intensity changes.

The crystals were strongly birefringent and were generally uncracked but some small cracks did exist on some of the larger crystals. The [001], [100] and [101] diffraction spectra were recorded on 9^{o} precession photographs and careful measurement of these gave the following cell dimensions:

$a = 56.1 \pm 0.4 Å$; $b = 66.7 \pm 0.5 Å$; $c = 81.3 \pm 0.5 Å$

Thus, the cell dimensions had changed by about $\frac{1}{2}$% except the c axis, which appeared barely to have changed. It might be, then, that the binding was at a different site from that in TCA.

Data Collection and Processing for the Projections.

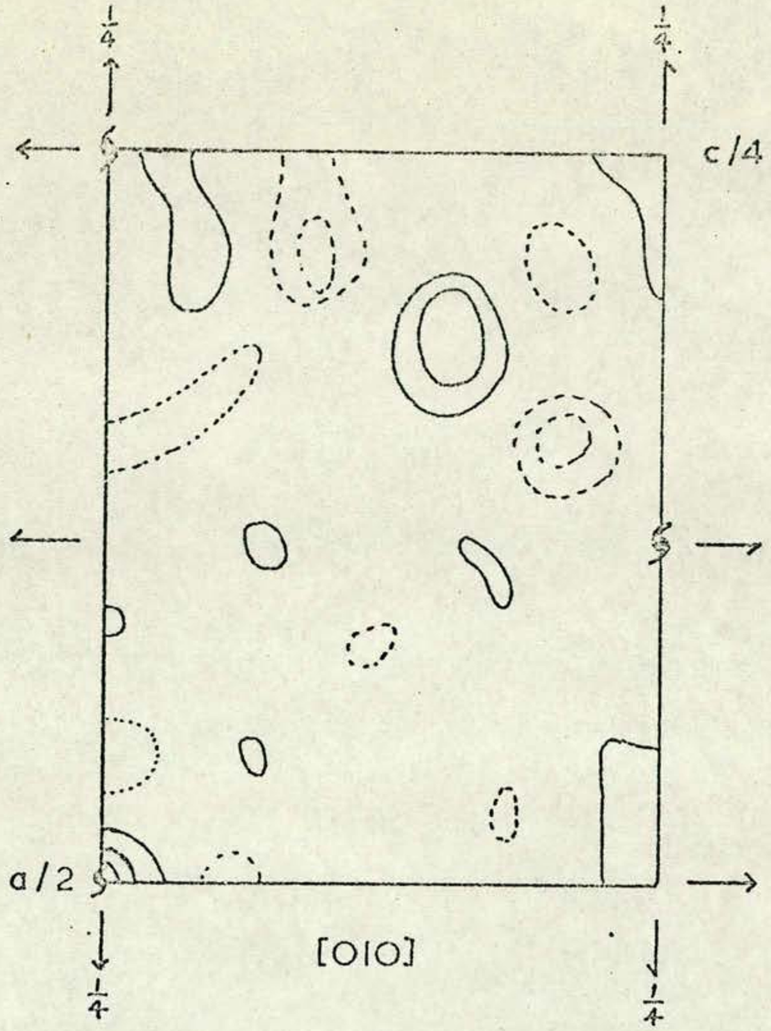
A fresh crystal was mounted with its b axis along the goniometer axis. This allowed 9^{o} photographs to be taken down each of the [100] and [001] directions. Because the crystal was large, the diffraction spectra were obtained in 9 hrs using the Phillips set as mentioned above. In order to obtain the third centric zone with the same crystal, the goniometer was transferred to the linear diffractometer. The setting was checked and corrected for the slight movement which had taken place during transfer. Once again, two cycles of the one minute motor were used to obtain an integrated peak count and the data were collected automatically over half of the zero layer. Having done this, the poor measurements were set by hand and remeasured so that the complete set was obtained. Finally, the absorption curve was collected on the 040 reflection, a conveniently strong, axial one.

The photographic data were corrected for Lorentz and polarisation effects as before and then compared with the native data in order to check if a fall-off correction was required. Although it appeared as though a small correction needed to be applied, there were too few terms to make the constants reliable and therefore no correction for this effect was applied at this stage.

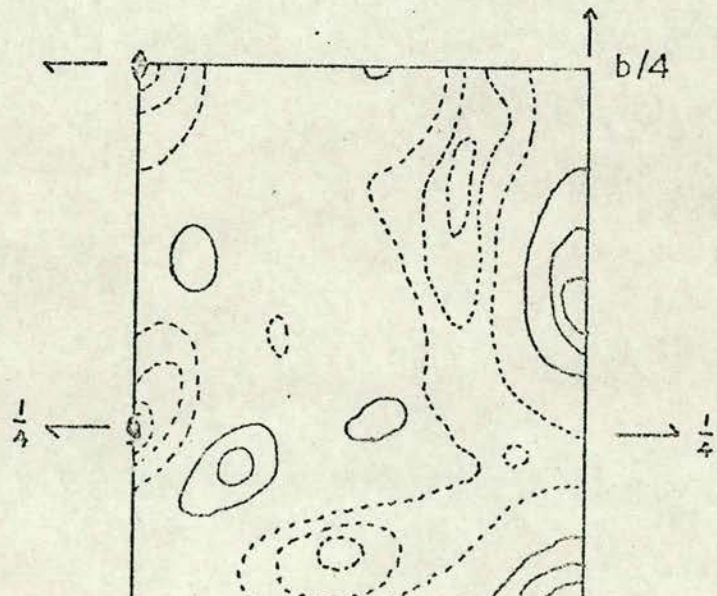
The diffractometric set of data was also corrected for Lorentz, polarisation and absorption effects as described for the TCA data. Because of a missetting of the limit switches, slightly more than the basic 6Å set had been collected. This enabled a determination of the corrections required for fall-off to be made. This was done, as for TCA, by comparison with the native set and the appropriate constants calculated. However, no correction was made at this stage since it was thought better to scale the data together and then to apply an overall correction to both the photographic and diffractometric data sets. A recalculation of the constants from all data would then show if the photographic set had a different fall-off. The two quarters of the zero layer which were measured by diffractometer, scaled together to give a value of 8% for D, defined as before. Scaling in the photographic data gave a final value of 10.2% for D for the difference between the two sets. The whole set was then compared with the native and constants for the fall-off correction obtained. These are shown in Table VII. 1. It was clear that the photographs had caused only a slight change. The fall-off correction was applied and the list compared once more with the native as a check.

Fourier Projections.

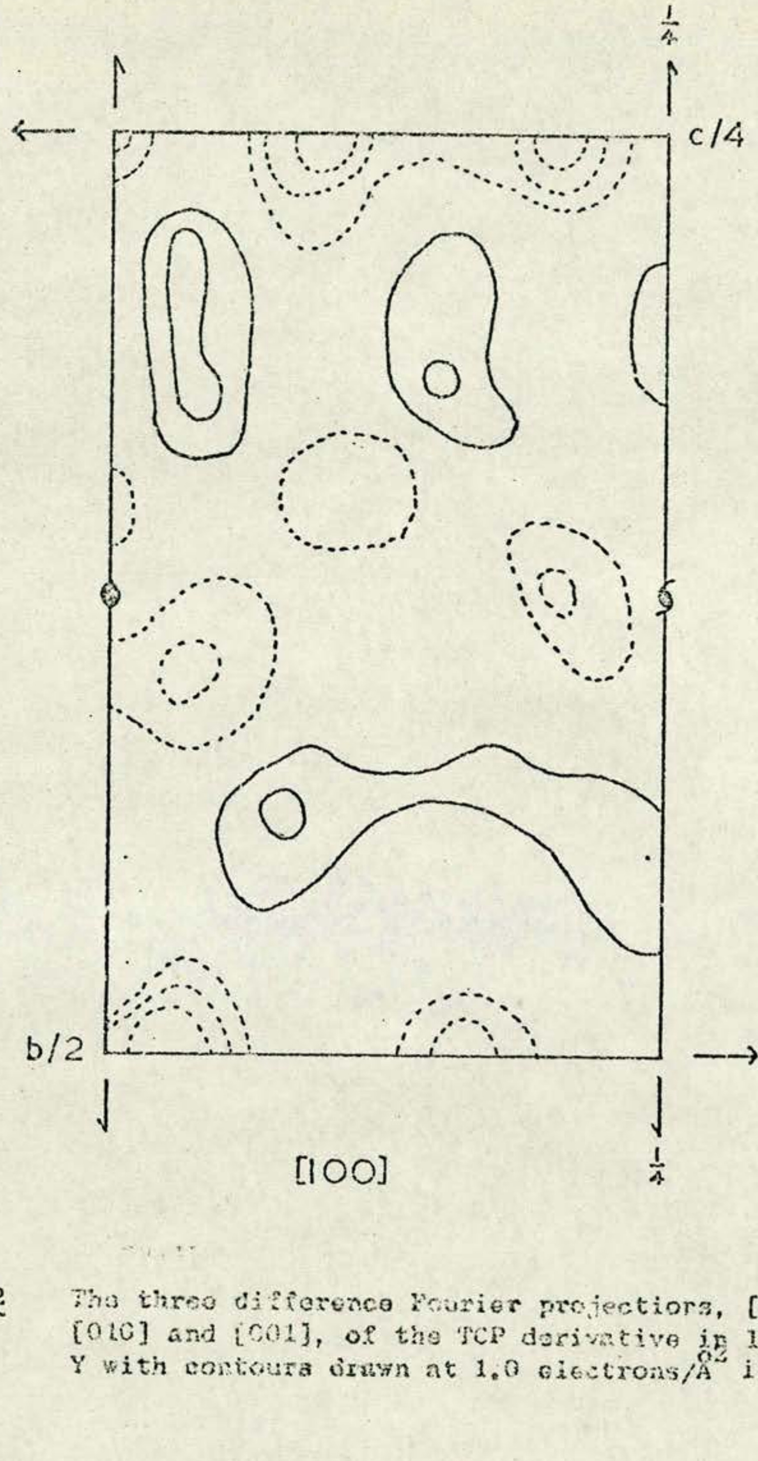
The three, principal difference Fourier projections were produced using the signs for the native set of data. These are



[010]



[001]



<u>FIG. VII.2</u> The three difference Fourier projections, [100], [010] and [001], of the TCP derivative in lattice Y with contours drawn at 1.0 electrons/Å^2 intervals.

shown in Fig. VII. 2, and it is obvious that, as before for TCA, the three projections have no single outstanding feature in common. However, the cleanest projection was that down the b axis and this showed a main peak corresponding approximately to the PTN site at x = 0.13 and z = 0.15. This feature was found to be common, more or less, to all three maps and was really the only positive peak. Two negative peaks were found to be common to all three, one of which was close to the MMA site indicating that there were probably more crossovers than normal, as had been found with TCA. Removal of some of the low order terms which had very large changes due to salt, had very little effect.

It was clear that the data would have to be collected in three dimensions in an attempt to overcome this problem. Before proceeding with this, however, it was decided to try a method suggested in a paper by Dickerson et al. (1967) whereby the refinement of an erroneous site will cause the occupancy of that site to plummet towards zero. It was hoped that PANGLOSS refinement would reveal any erroneous site chosen and that, using the calculated signs obtained after completion of the refinement, a double-difference Fourier would show up any minor sites.

Choice and Refinement of Trial Heavy Atom Sites.

There were three obvious trial sites. The MMA site could be occupied because in Lattice Y the isoelectronic TCA binds there. The HGI site was also occupied in the TCA derivative and must therefore be considered a possibility. Finally, $Pt(NO_2)_4^=$ is a square planar molecule and, whilst the large NO_2^- groups might be responsible for the binding in the third site, it was thought

HGI

		15 Cycles	7 Cycles	10 Cycles
Scale	1.000	1.000	-	0.82
Occupancy	51	68	-	65
B	20	-	16.75	-
x	0.056	0.054	-	0.054
y	0.964	0.961	-	0.961
z	0.230	0.236	-	0.238
R.M.S.	204	191	191	180
R	0.642	0.595	0.596	0.586

MMA

		15 Cycles
Scale	0.857	-
Occupancy	50	60
B	16.73	-
x	0.239	0.234
y	0.039	0.038
z	0.250	0.261
R.M.S.	202	183
R	0.705	0.616

PTN

		10 Cycles
Scale	0.857	-
Occupancy	50	60
B	16.73	-
x	0.133	0.146
y	-0.163	-0.164
z	0.199	0.201
R.M.S.	189	178
R	0.618	0.580

TABLE VII.2 Trial refinements of the three estimated sites in the TCP derivative. The sites have been taken singly.

	Initial	16 Cycles
Scale	0.857	0.807
Occupancy	59	47
B	16.73	-
x	0.14575	0.14762
y	0.8360	0.8365
z	0.20139	0.20339
Occupancy	60	69
B	16.73	-
x	0.234	0.23337
y	0.03844	0.04344
z	0.26106	0.261
Occupancy	65	40
B	16.73	-
x	0.054	0.052
y	0.96137	0.96287
z	0.23786	0.23836
R.M.S Del	225	194
R	0.533	0.507

185 Reflections.

TABLE VII.3 Refinement of the three trial sites simultaneously for the TCP derivative.

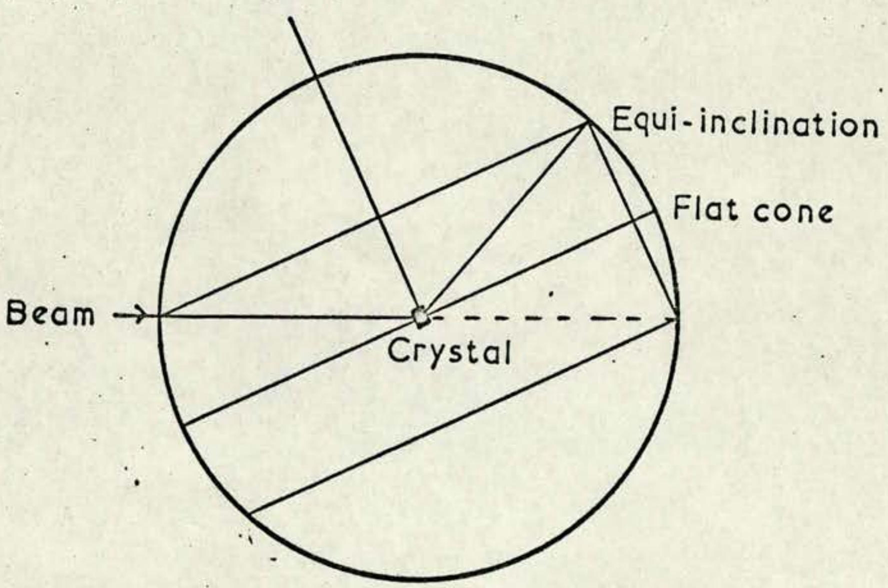


FIG. VII.3 Section through the sphere of reflection showing the flat-cone setting for data collection.

possible that a similarly charged complex of the same general shape and nucleus might also tend to associate at the PTN site.

Consequently, the parameters for each of these three sites were fed into PANGLOSS in turn and about 12 cycles of refinement carried out. The scattering curve was that of a platinum atom and the temperature factor was not refined. The results of the three refinements are shown in Table VII. 2. The results were rather discouraging because all three sites appeared to have approximately equal occupancy, a situation which, although quite possible was undesirable for further work with this derivative. In order to check that each occupancy was not artificially high, all three sites were refined together. These data are found in Table VII. 3. Adding in all three sites gives an r.m.s. error between observed and calculated F_H of some 7% higher than that obtained in any of the single refinements but the R-factor, a more sensitive criterion, had decreased by about 15%. On the whole, therefore, slight preference seemed to be towards the MMA site followed by the PTN with the HGI last, but no clear cut result had been obtained. Really the only way to clarify the situation and determine whether the sites obtained in projection were, in fact, genuine was to collect three dimensional data. This was done in the hope that one or possibly two sites might show up clearly, as had happened with TCA.

Three Dimensional Data Collection.

The diffraction spectrum of lattice Y, as has been said, has only reflections for which $h+l = 2n$. This means that, if the b

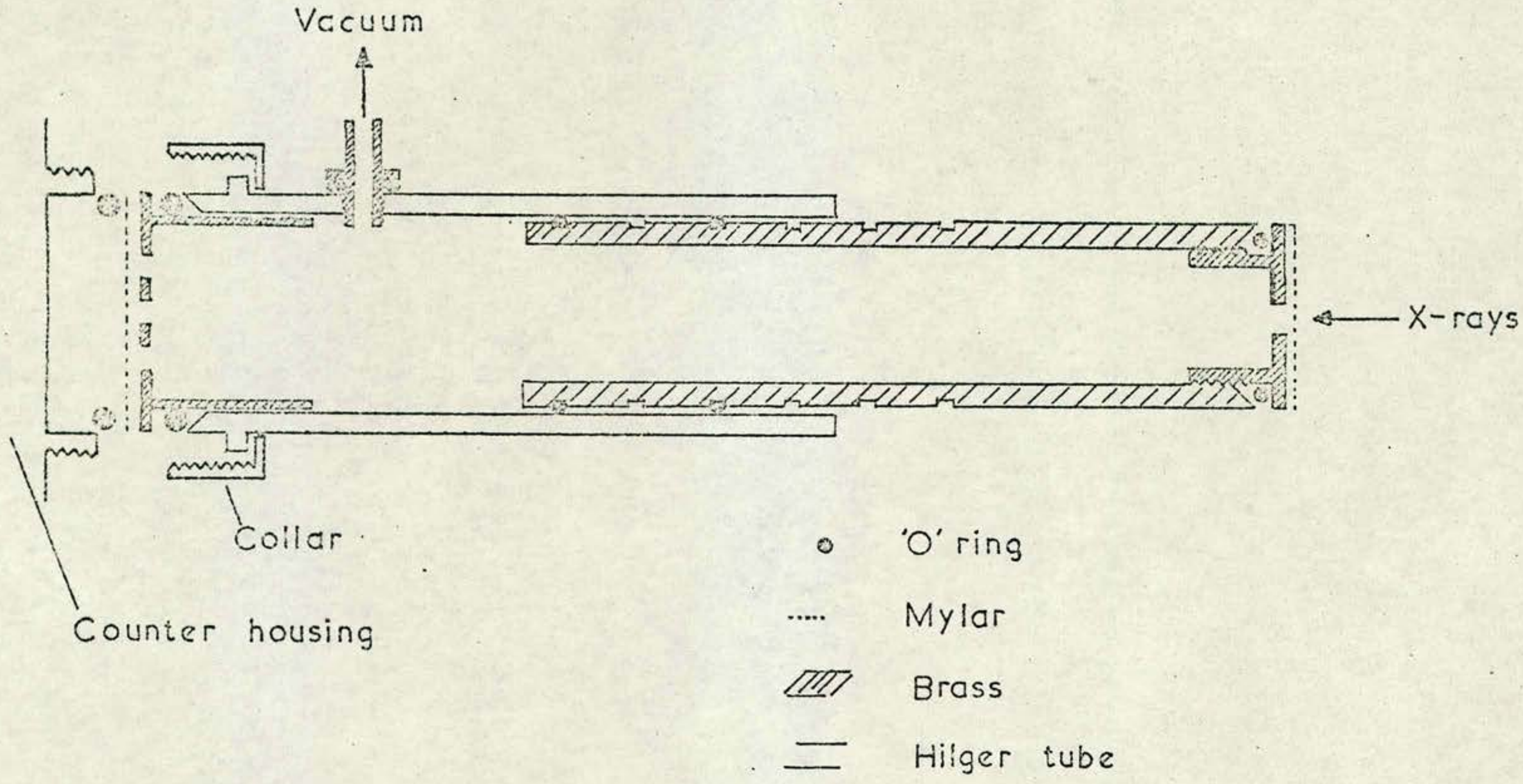
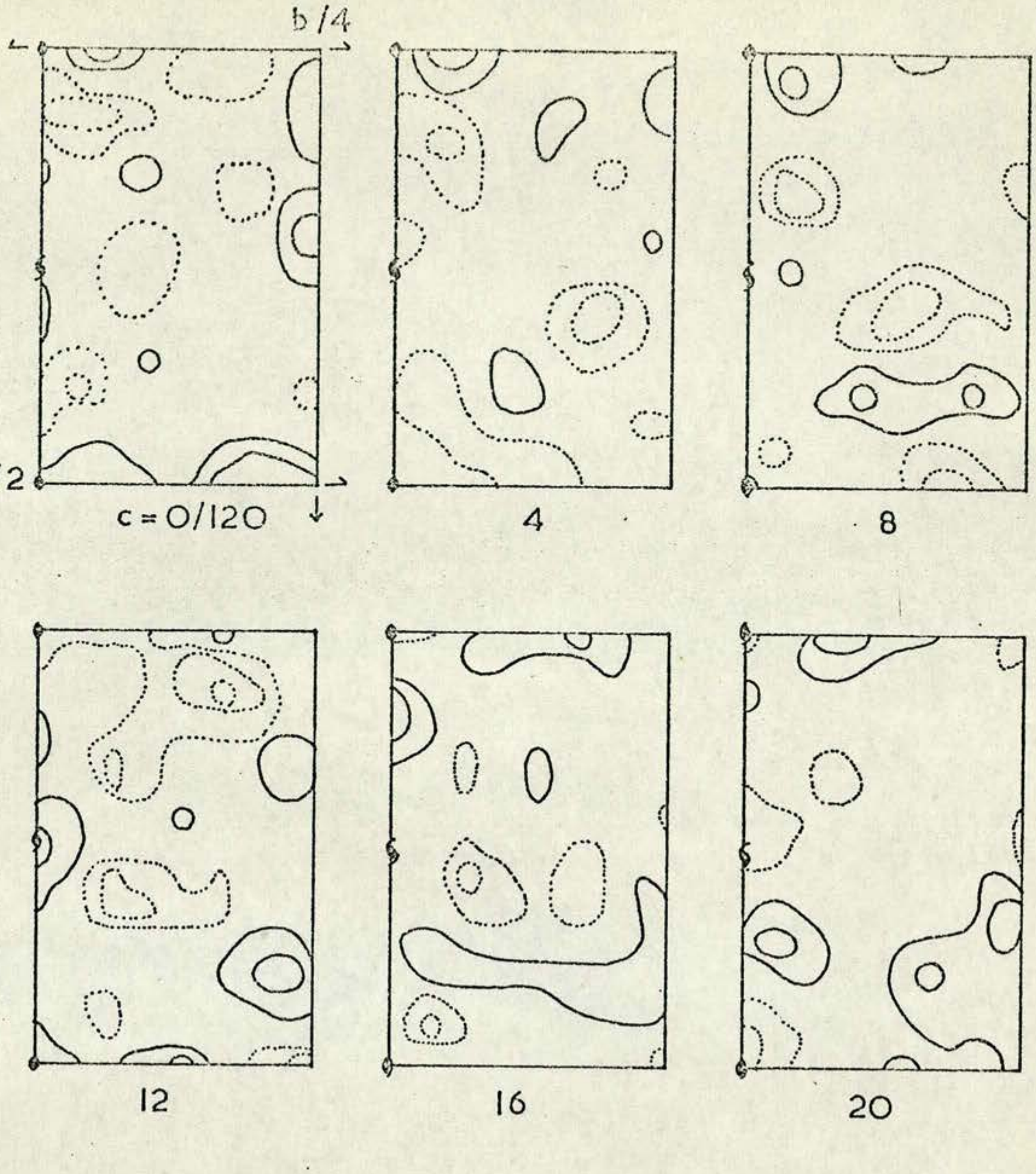


FIG. VII.4 Variable secondary collimator for the linear diffractometer allowing three-counter operation and also evacuation of the beam path. The diagram is not drawn to scale.

axis is mounted parallel with the goniometer axis of the linear diffractometer then points of equal h and l but differing k will lie directly above one another. When the unit cell is sufficiently large, as in the case of proteins, it is possible to satisfy Bragg's Law quasi-simultaneously for three reflections as shown in Fig. VII. 3. The adaptation of the linear diffractometer to measure three reflections at once has been described by Arndt, North and Phillips (1964) and such an instrument was available in the laboratory. The matched proportional counters were fixed 0.7cms apart so that an extension tube had to be fitted to give a crystal to counter distance such that this separation occurred between neighbouring or alternate reciprocal lattice layers. In lattice Y this distance was 32.8 cms for adjacent layers and, because this was so long, it was thought advisable to evacuate the counter extension tube in order to increase the number of counts obtained in a given time.

An extendable tube was designed and built for the purpose. With thin mylar windows and a pressure below 60mm Hg, an increase of 30% in the number of counts detected in a minute was obtained. The design is shown in Fig. VII. 4. The problem of attaching the end windows to form as good a vacuum as possible was overcome by Mr. D. Pettigrew. He found that whilst glue was sufficient at the beam entrance where the hole was small compared to the area available for adhesion, it failed to give satisfactory results at the counter end. At this end, therefore, a thin film of grease was applied to the end plate of the tube, the mylar stretched over this and then clamped in place by an O-ring between the counter


b/4
a/2
c = 0/120
4
8
12
16
20
24
28
32

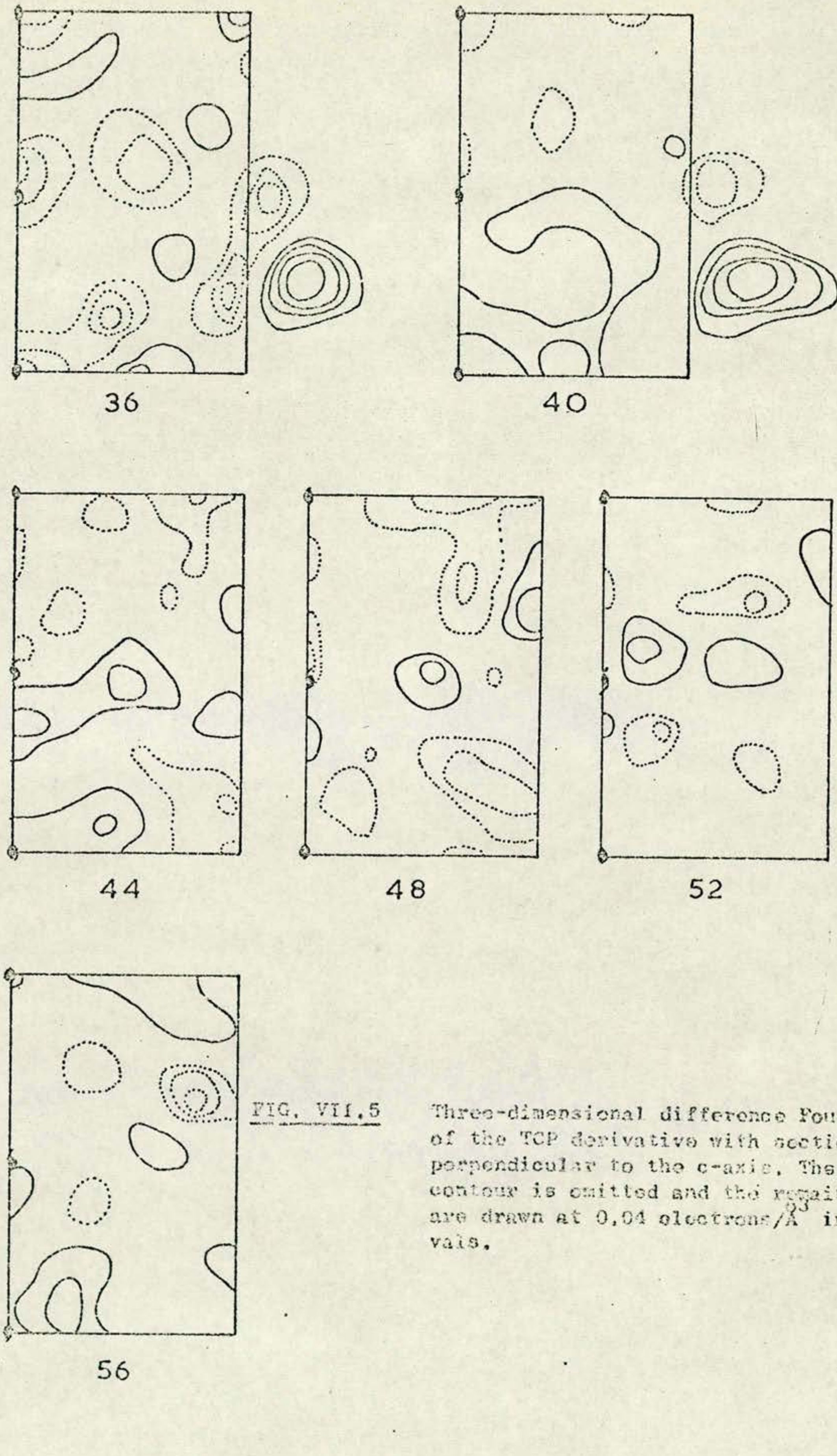


FIG. VII.5 Three-dimensional difference Fourier of the TCP derivative with sections perpendicular to the c-axis. The zero contour is omitted and the remainder are drawn at 0.04 electrons/Å^3 intervals.

	Initial	12 Cycles	9 Cycles	4 Cycles	5 Cycles	10 Cycles
Scale	0.750	0.735	-	0.749	-	-
Occupancy	54	47	-	-	-	47
B	20.000	-	-	-	20.006	-
x	0.344	0.336	-	-	-	0.341
y	0.298	0.308	-	-	-	0.310
z	0.363	0.353	-	-	-	0.352
Occupancy	35	--	35	-	-	35
B	20.000	-	-	-	19.906	-
x	0.999	-	0.981	-	-	0.981
y	0.061	-	0.043	-	-	0.043
z	0.019	-	0.025	-	-	0.025
Occupancy	48	53	-	-	-	53
B	20.000	-	-	-	19.931	-
x	0.050	0.043	-	-	-	0.040
y	0.050	0.047	-	-	-	0.050
z	0.240	0.241	-	-	-	0.241
R.M.S. Del	203	185	173	173	174	172
R	0.580	0.534	0.511	0.512	0.512	0,501

139 Reflections.

Zone	Number	R.m.s. d*	$\overline{F}_p$	$\overline{F}_{ph}$	$\overline{F}^{obs}_h$	$\overline{F}^{calc}_h$	$\overline{Del}$	$\overline{R}$
0	15	0.052	418	367	288	289	143	0.494
1	23	0.081	526	350	312	246	129	0.519
2	22	0.105	582	490	192	218	100	0.458
3	25	0.124	538	379	276	253	147	0.580
4	22	0.142	391	331	200	136	133	1.012
5	23	0.155	375	283	286	250	141	0.565
6	9	0.164	394	319	361	258	139	0.541

TABLE VII.4 Refinement of the centric data for the TCP-lactoglobulin derivative in lattice Y.

housing and the tube,secured by the counter retaining collar. This gave a system which required only occasional re-evacuating.

As before, only a 6Å data set was collected over an octant of reciprocal space. This set was corrected for Lorentz, polarisation and absorption effects and any unsatisfactory measurements were repeated. In this way, the near complete set was obtained. Because the flat cone setting left out some reflections near to the vertical, b axis, there are slightly more unmeasured reflections than for TCA. The layers were scaled to the native in order to obtain the values for fall-off correction and also because there were far less common reflections had the photographic set been used. Values for A and B for the individual layers were not used on account of the small number of terms with which they would have had to be calculated. Instead average values were obtained for all layers together. The values of the A and B used,along with D, the comparison between this set and the previous one, are shown below.

$A = 1.205 \quad ; \quad B = -11.905 \quad ; \quad D = 11.71\ \%$

The final, corrected data set for TCP is shown in Appendix III. This set was used, along with the data for the native, to produce a three-dimensional difference Fourier which is shown in Fig. VII. 5. Straightaway, it can be seen to be very much more "messy" than that for the TCA derivative. The three largest peaks were taken and fed to PANGLOSS. The refinement of these is shown in Table VII. 4 and it must be regarded as rather unsatisfactory because only the centric data were used. Nevertheless, the 'major' site in the Fourier map appeared to be in the same region as the PTN one, being some 5.8Å distant, mostly in the z-direction. It had, however, a lower occupancy on refinement than the larger of the two minor sites in the Fourier but it must be remembered that the refinement only used about one third of the data. Of the other two sites, both were new and the one close to 0 0 $\frac{1}{4}$ seemed to have appreciable occupancy. No real indication of occupancy at either the HGI or the MMA sites was obtained.

A double difference Fourier projection using calculated signs showed there to be two small peaks close to those already found as minor peaks in the difference Fourier above. This indicated that the observed data was not being well enough fitted by the calculated set. The scarcity of data in the centric zones was making the use of PANGLOSS inefficient and a full three dimensional refinement should really have been carried out in order to obtain accurate coordinates and occupancies. This unfortunately could not be done because the three dimensional data for the MMA, PTN and HGI derivatives were not available on disc and therefore recalculation of the phases, a necessary intermediate between two refinement cycles, could not be done.

One thing seemed clear and that was that no useful phase information was going to be obtained from the TCP derivative unless all of the sites could be identified and refined. However, before abandoning TCP completely and when the data for the various derivatives become available on disc, it would be worth trying a three dimensional refinement in the hope that clarification might be obtained of the situation as regards the occupancies of the various sites and, indeed, whether any other undetected sites exist.

CHAPTER VIII

Further Studies

Introduction.

In the previous chapter, mention was made of the inapplicability of u/v spectrophotometry to the TCP-β-lactoglobulin complex because of the absorption maximum's coming on top of the region of protein absorption. This, however, did not affect the polarimetric measurements but any tie-up between the solution and X-ray work became less certain since no indication could be obtained as to where the binding was occurring, save that it did not appear to involve the free cysteine. It was decided, therefore, to examine a series of related compounds in the hope that (a) some idea of whether the size of the region in which the cysteine was situated was an important factor in the binding of ligands to it, (b) whether other changes in the spectrum and specific rotation could be correlated to a specific reaction to a specific side-chain of the protein other than the sulphydryl, and (c) whether the overall charge on the ionic species which was binding made a noticeable difference to the site at which the binding was occurring. From the TCP work it appeared that this latter effect was of importance since the overall negative charge had increased from that of TCA and no binding at the SH seemed to be occurring in solution. On the other hand, might the oxidation state be the controlling factor, acting as a "driving force" for the reaction of TCA and, if this were so, how could it best be examined?

The other oxidation state in which platinum can exist (other than O) is IV. By examining the reaction with $Pt(CN)_6^{=}$, an octahedral complex, it was hoped that it might be possible to find out if the sulphydryl could start to reduce the Pt(IV) as it had the Au(III). The reaction, of course, was expected to be slower because of the larger size of the octahedral complex and the consequent steric hindrance.

The larger size of the $Pt(CN)_6^{=}$ and the stronger binding of the CN^- ions might mean that the reduction/reaction might be stopped completely. In case this happened, an examination of $PtCl_4^{=}$ and $PtCl_6^{=}$ was also planned. Finally, moving to iridium, the next element of lower atomic number to platinum, the readily obtainable chloride complexes are those of oxidation states III and IV. Both of these are octahedral so that the reduction of the Ir(IV) complex would not give a smaller,and hence possibly more easily accommodated, complex as in the case of Pt(IV) and Pt(II).

Some preliminary diffusion trials with these complexes had been carried out in lattice Y (Green, unpublished work), the results of which will be discussed below. However, lattice Y was thought to be not such a good form for high resolution work because of the proximity of the heavy atoms' z-coordinates to $\frac{1}{4}$. Of the two high pH forms, therefore, lattice Z looked the more promising and it was with this form that the diffusion and crystallographic work was done.

<u>Hexacyanoplatinate (IV)</u>

Information in the literature on the u/v spectrum and structure

FIG. VIII.1 Ultraviolet spectra of K_2PtCl_6 and $K_2Pt(CN)_6$ (?).

of this ion was found to be singularly lacking. If it is assumed that the Pt-C and C-N bonds are roughly the same length as they are in the TCP complex, then the overall shape of the ion is roughly spherical with a radius of about 3.5Å. This is a much bulkier molecule than either of the more disk-like TCA or TCP ones, so that any reaction with the -SH group, whether substitution or complete reduction, was likely to be much slower.

No stability constants for the hexacyano complex could be found and only K_6 for the hexachloro one. In order to prepare the $Pt(CN)_6^{=}$ complex it was assumed that the CN^- ions would replace the Cl^- ones in the more readily available $PtCl_6^{=}$. This was because a typical order of decreasing equilibrium constants is (Basolo and Pearson (1967))

$$CN^- > NH_3 > OH^- > I^- > Br^- > Cl^-$$

Although it was realised that complete replacement was unlikely, it was hoped that the greater affinity of cyano over chloro-ions and the excess of the former ion used would force the equilibrium to favour the $Pt(CN)_6^{=}$ species. The spectrum of the solution formed by treatment of K_2PtCl_6 with a 9:1 excess of KCN is shown in Fig. VII. 1 together with that of pure $PtCl_6$. These show distinct differences,indicating that at least some cyanide has been taken up but, since no spectrum was available, it was not sure whether this was, in fact, the required hexacyanoplatinate or merely a mixture of all possibilities.

This problem is not restricted to this particular case, of course. The relative stabilities of the complexes used in the

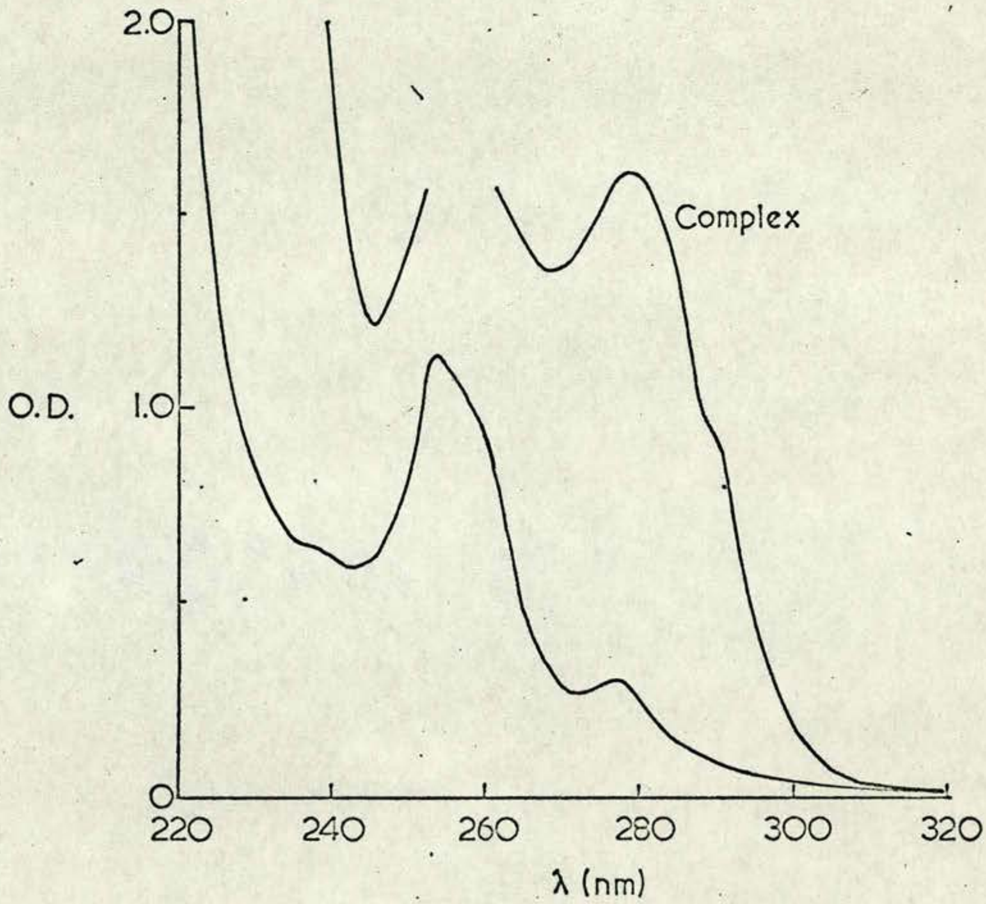


FIG. VIII.2 Difference spectrum of the $Pt(CN)_6^{=}$-protein complex versus $Pt(CN)_6^{=}$ at pH 7.8. There was a 13:1 excess of platinum complex and the protein was 1.001×10^{-4} M.

isomorphous replacement method with respect to substitution by other species present in strong salt solution is a very complicated one. Fortunately the most likely substitutions to occur are those involving NH_3, H_2O and possibly OH^- which do not introduce large groups or heavy atoms capable of upsetting the location of the heavy atom. They do, however, introduce the possibility of a change in the overall charge of the complex which may cause quite different binding from that expected. The answer to this requires both more detailed chemical information on the existence in various solutions of the complexes used for the isomorphous replacement method and also the correlation of this information with the specific binding of these complexes to the actual proteins themselves. At the present time, not enough high resolution work on the binding of heavy atom complexes to proteins is available although much work is being done on the binding of amino acids and peptides to heavy atoms (see, for example, the review by Freeman (1967)).

The spectrum of $Pt(CN)_6^=$ shows that a peak occurs at 254nm which, once again, is in the region of the protein minimum but, in this case difference spectra were possible since the molar extinction coefficient of the peak was found to be 2200 cm^2/mole. The spectra of the β-lactoglobulin-$Pt(CN)_6^=$ complex and the difference between that and β-lactoglobulin are shown in Fig. VIII. 2. The concentration of the protein was 1.001×10^{-4} and the platinum complex was present as a 13:1 excess. If the $Pt(CN)_6^=$ peak at 254nm is subtracted from that obtained with the protein a residual peak at this wavelength is obtained. This was an indication that some reaction was occurring between the protein and the hexacyanoplatinate in solution. Unfortunately, however, this change required a very long time to reach completion (about 10 days) at a pH of 8.45. The higher-than-expected pH was probably because of the excess CN^-, some possibly released from the complex, removing hydrogen ions.

This change of u/v spectrum was paralleled by an equally slow change in conformation to the value associated with the reaction with TCA. However, because the time required for reaction was so long, it was not inconceivable that the conformational change was merely the slow denaturation of the protein in the presence of heavy metal ions (cf. Pantaloni (1965)).

Both of the above indications could be taken as showing that the $Pt(CN)_6^{=}$ could react, possibly by a similar mechanism of substitution (other than reduction) as was found with TCA, with β-lactoglobulin. Whether the reaction was with the sulphydryl group was not determined, but by analogy it could have been. The slowness of the reaction and the possible existence of other species as well as the hexacyanoplatinate made the whole solution study rather unsatisfactory.

From the X-ray point of view, the following information was available from Lattice Y studies by Green and coworkers (unpublished). A 2:1 heavy atom to protein ratio was found to give no changes at all after about three month's diffusion. A further trial with an increased ratio gave large changes but also caused a greater amount of fall-off than for the native and also, the lattice was distorted in such a way that the b axis had increased to 68.2Å.

The a and c axes did not appear to have changed significantly. This last trial was made with $PtI_6^=$ instead of $PtCl_6^=$ as the platinum complex, and the cell dimensions correspond exactly to those of β-lactoglobulin diffused with K_2PtI_6 alone. It could be, therefore that the CN^- ions had not replaced the I^- ones, an observation supported by the sequence shown on page 151

The $Pt(CN)_6^=$ stock solution made up for the experiments reported above contained only a tiny amount of insoluble material compared with that originally present, i.e. K_2PtCl_6. This was taken as a possible indication of the greater solubility of $K_2Pt(CN)_6$. However, a slightly more soluble form of the $PtCl_6^=$ species is the ammonium salt (Chemical Rubber Handbook) and consequently attempts to diffuse the hexacyano complex into lattice Z were thought to be more likely in concentrated ammonium phosphate buffer than in the mixed sodium/potassium phosphate buffer from which the crystals were grown, since the cyanide ions would have a better chance to replace the chloride ones if the complex was in solution. Attempts were made in both sorts of buffer at concentrations of 2mM, 10mM and 20mM. The salt concentration was 2.2M in phosphate, the pH was between 7.8 and 8.0, there was an excess of KCN of about $1\frac{1}{2}$:1 for each CN^- required and the crystals were taken from a tube containing small but well-formed, highly birefringent lattice Z crystals which had been prepared by Green and Wishnia (unpublished).

13^o (hol) precession photographs of the most concentrated samples were taken about four weeks after the diffusion had been started. These showed there were only the smallest changes at medium resolution although there were fairly large low resolution changes, caused by the change in mother liquor. After three months in the same solution, the crystals, though outwardly in good condition, had little birefringence and were very poor diffractors of X-rays. A 30-hour exposure was required to give but a feeble diffraction spectrum, although it did seem to extend to about 3Å.

Because of the slowness of reaction and the lack of any useful effect on the soaked crystals, the hexacyanoplatinate experiments were stopped. Recently, a new preparation of solid $K_2Pt(CN)_6$ has been published (Babkov (1967)) and, with a pure sample of this material, it might be worth repeating the above solution and diffusion work in order to check that the results obtained were, in fact, due to $Pt(CN)_6^{=}$.

Hexachloroplatinate (IV) and Tetrachloroplatinite (II).

Two reasons were responsible for the examination of the tetra- and hexa-chloro complexes of platinum. First, the solution work on hexacyanoplatinite had pointed to the possibility of the free sulphydryl groups ability to reduce, or, at any rate, to react covalently with the heavy atom complex in a similar manner to the interaction with TCA. The reaction was, however, slow and some means of speeding it were sought. By using a less tightly bound and slightly smaller complex, some idea of the stoichiometry of the interaction (if, indeed, there proved to be any) might be obtained without running the risk of bacterial contamination or denaturation whilst the solutions were standing to equilibrate.

Second, the tetrachloro complex has been found to be specific for the methionine sulphur atom if this is accessible (Dickerson et al. (1969)). It also reacts with free sulphydryl groups, presumably by a replacement-type mechanism of the sort detailed by Cattalini et al. (1967). Dickerson also noticed that in cytochrome c there was evidence of oxidation to Pt(IV), the extra ligands coming from the methionine sulphide and the mother liquor. This was thought to be of sufficient interest to warrant investigation especially since the evidence from the TCA was that there was a reduction.

Ginstrup and Leden (1968) have found the oxidation potentials of the system to be those shown below.

$$PtCl_6^{=} \xrightarrow{+0.529v} PtCl_4^{=} \xrightarrow{+0.547v} Pt$$
$$PtCl_6^{=} \xrightarrow{+0.551v} Pt$$

These would appear to favour a reduction of the platinum at the expense of the cysteine. Initial trials with 1.02×10^{-4}M protein, 2.72×10^{-4}M $PtCl_6^{=}$ at pH 7.66 showed there to be a reaction which went on past the intermediate state found with TCA and seemed to give rise to increasing quantities of denatured protein as though the R state was being pushed right through to the S state as had been found by Pantaloni (1965) with Hg^{++} ions. This effect, however, was much slower, taking a week for an increase of about 15^{o} in the specific rotation (589nm), corresponding to about half that required to give the denatured protein's specific rotation.

What was happening? Was the $PtCl_6^{=}$ binding as $PtCl_6^{=}$ to two sites which, together, were capable of causing denaturation or

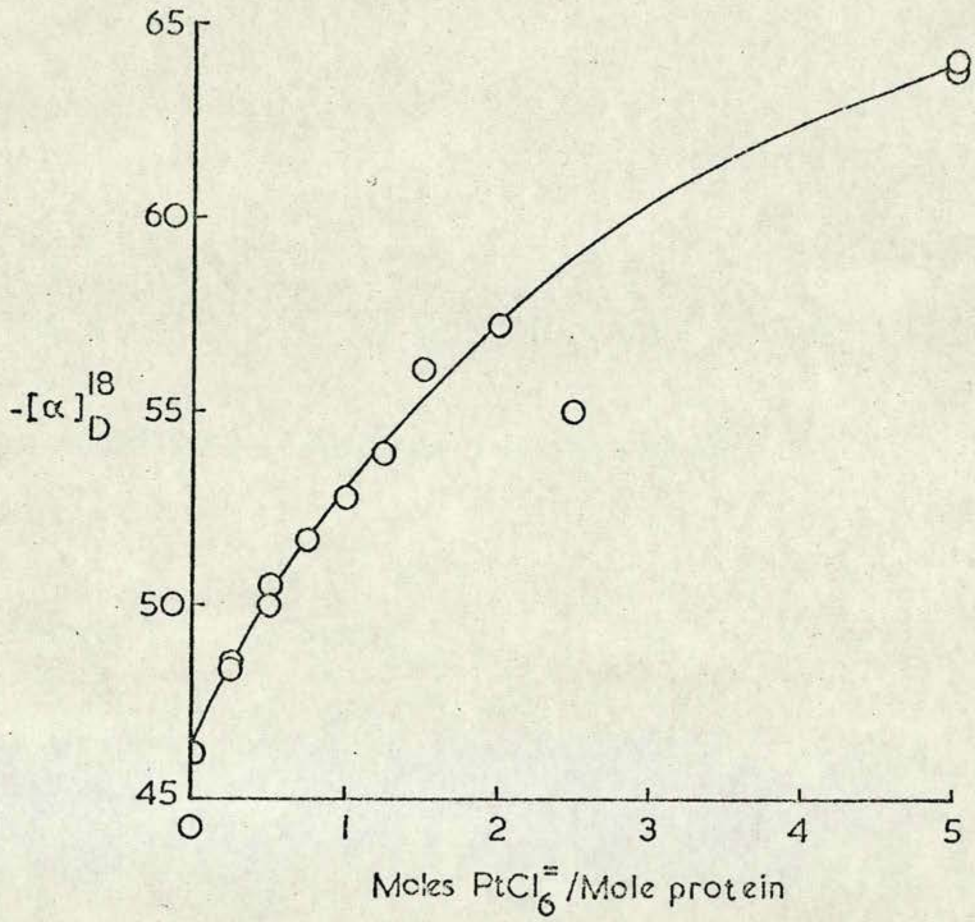


FIG. VIII.3 Specific rotation of the $PtCl_6^{=}$-lactoglobulin complex as a function of the number of moles of platinum added per mole of 1.02×10^{-5}M protein at pH 7.8.

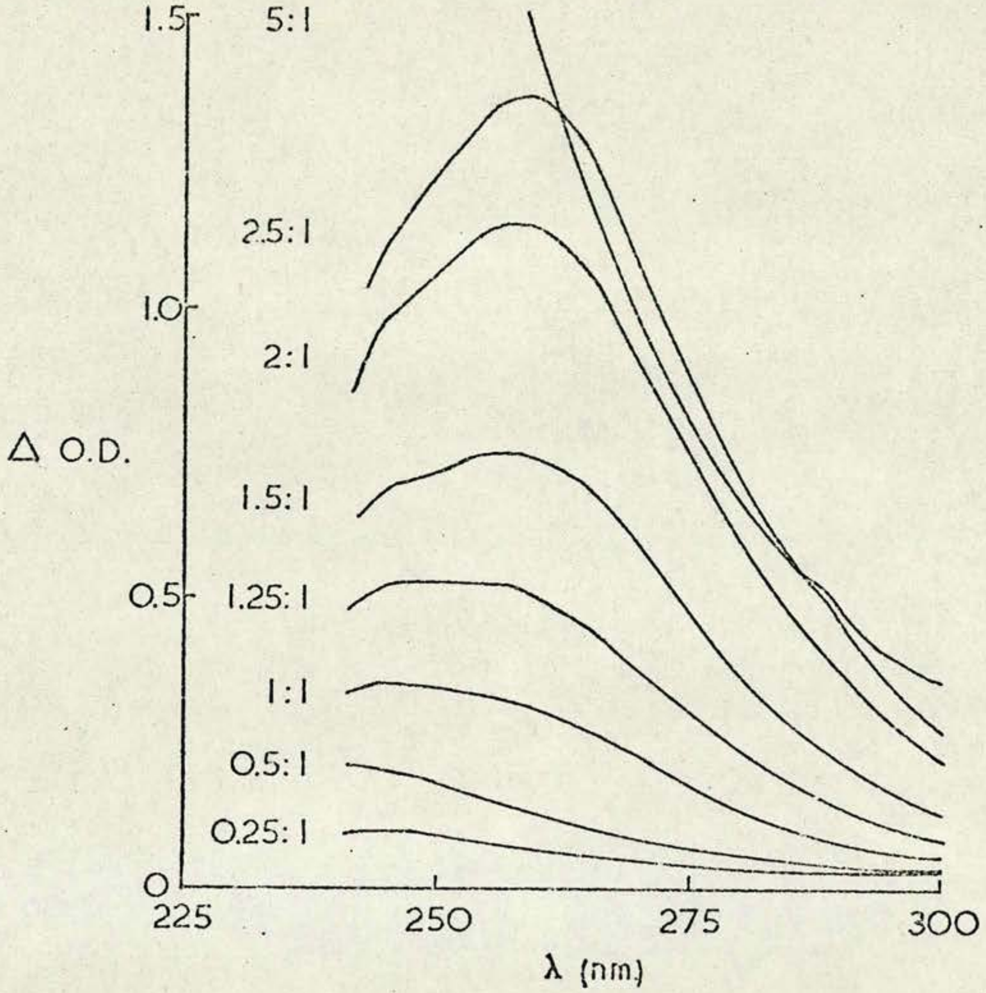


FIG. VIII.4 Difference spectra of the solutions from Fig. VIII.3 recorded on the SP800.

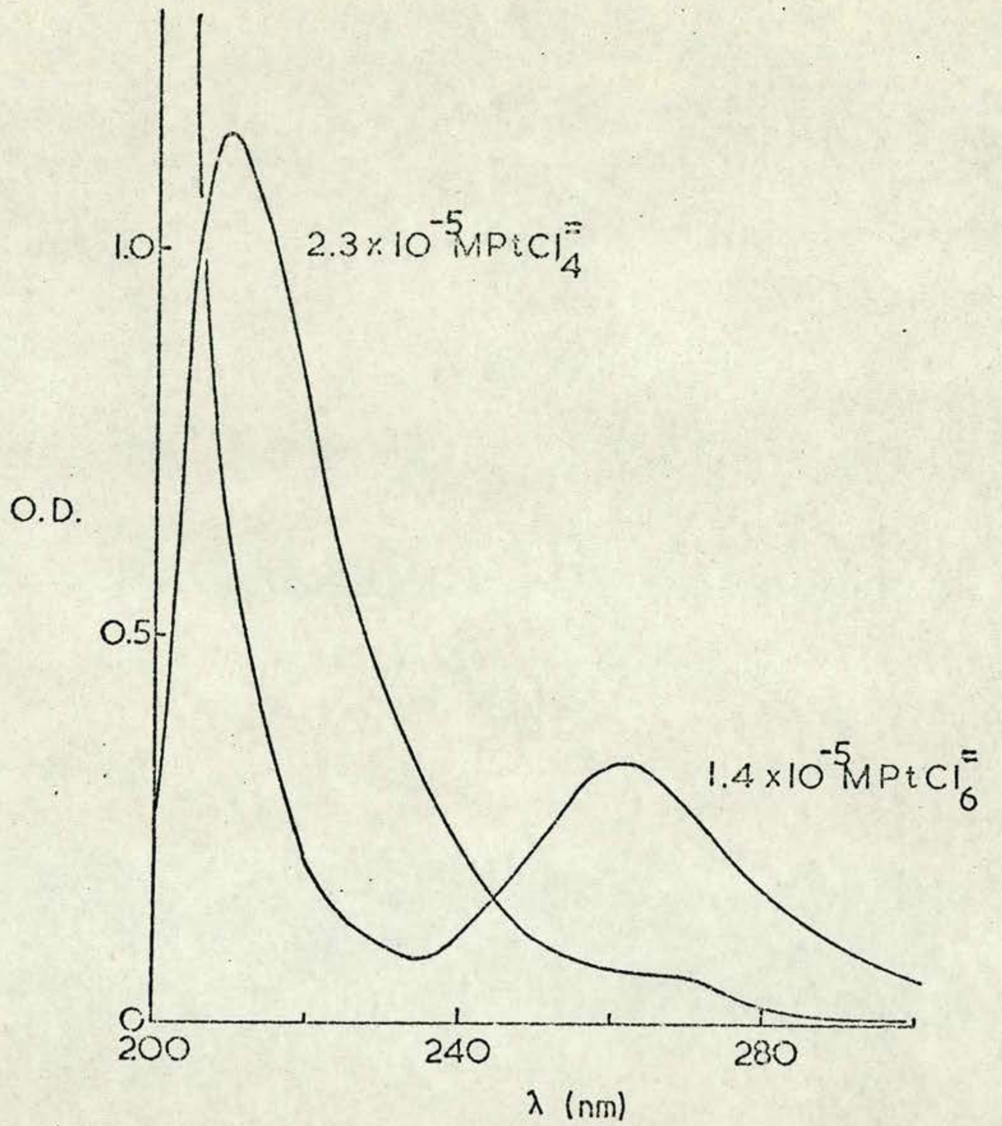


FIG. VIII.5 Spectra of 1.4×10^{-5}M $PtCl_6^{=}$ and 2.3×10^{-5}M $PtCl_4^{=}$ solutions in 0.1M phosphate buffer at pH 7.8.

was the $PtCl_6^{=}$ being reduced to $PtCl_4^{=}$ and this, when liberated, attacking some other group thus causing the denaturation? Hydrolysis of $PtCl_6^{=}$ is a common occurrence (e.g. Davidson and Jameson (1965)) and this too could have a bearing on the problem by altering the overall charge of the binding species. The first stage in an attempt to clarify the situation was to determine the number of moles of heavy atom binding per mole of protein. Accordingly, a series of preparations were made up with varying amounts of $PtCl_6^{=}$ stock solution, added as thoroughly shaken suspension, because the solubility of the K_2PtCl_6 is about 5g./1. at 20°C. The extra dilution, however, allowed all of the suspension to dissolve. The reaction was left to equilibrate at 3°C for $24\pm\frac{1}{2}$ hours and then the optical rotation measured and difference spectra run on the SP800. The optical rotation of the solutions increased slightly over the next 24 hours but further measurement was rendered inaccurate because filtration was necessary to achieve stable polarimeter readings, the precipitate being one of denatured protein and consequently the concentration was not precise enough to give reliable values for the specific rotation. The data from both the initial optical rotation and difference spectral measurements are shown in Figs. VIII. 3 and 4. The spectrum of a similarly diluted sample of the stock $PtCl_6^{=}$ solution in phosphate buffer is shown in Fig. VIII. 5. The molar extinction coefficient at 262nm is 24500 from which the concentration was calculated as 1.4×10^{-5} M which agrees within experimental error with the amount which should have been present.

Once again the strong protein absorption band from about 240nm to shorter wavelengths makes the difference spectra uninterpretable below about 245nm. It would appear, however, that the peak at 262nm moves to shorter wavelengths and in fact broadens at complex to protein ratios of less than about 0.7:1. The optical rotation data also showed a gradual increase above a 1:1 ratio and nothing like the precise cut-off obtained with TCA. Coupled to this increase above a 1:1 ratio was the change of the specific rotation measurements with time. Binding was occurring, apparently, at more than one site and in such a way as to cause denaturation.

Because of the ease of hydrolysis of $PtCl_6^{=}$, it was possible that not only was there reaction at the sulphydryl site but also at other sites such as methionine and histidine and cystine and that the binding was of a much more definite type than mere "association" by van der Waal's-type forces. This will only be found out by examination of the binding sites of the various species once the full structure is known.

The X-ray data from lattice Y was as follows. At complex to protein ratios up to 1:1 there were almost no changes, above 1:1 only slight changes were apparent until a ratio of at least 3:1 at which point large changes and a high fall-off occur. The cell dimensions also change to the values shown below.

	a(Å)	b(Å)	c(Å)
Native	55.6 ± 0.3	67.2 ± 0.4	81.7 ± 0.4
Derivative	55.7 ± 0.3	67.7 ± 0.4	82.8 ± 0.4

Thus, once again the quantity of reagent seems to be important and

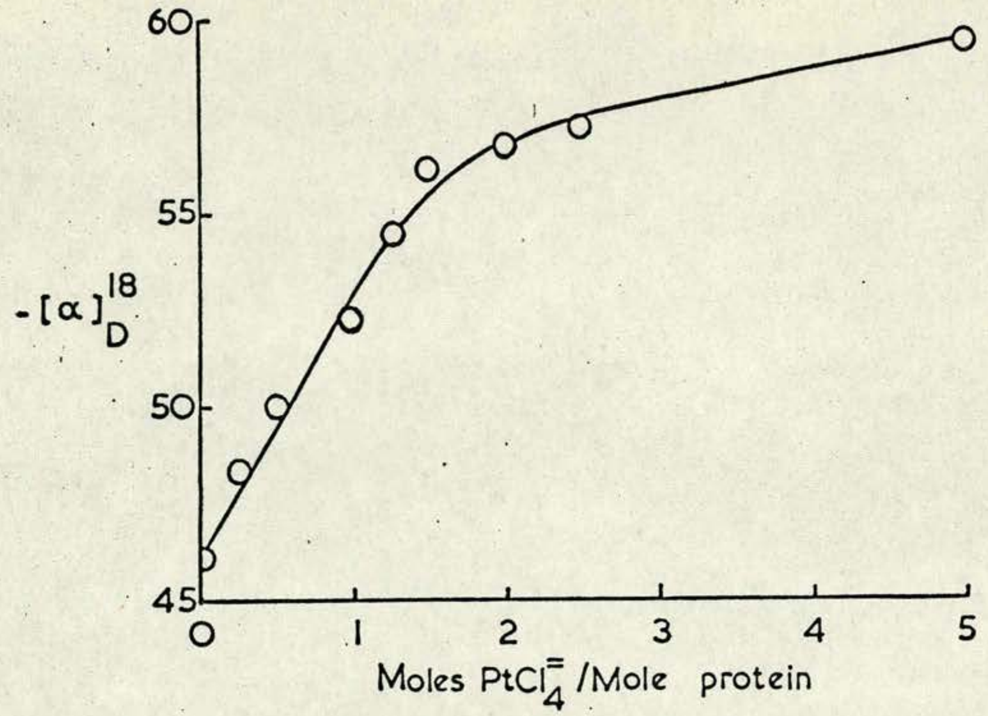


FIG. VIII.6a Specific rotation of the $PtCl_4^=$-lactoglobulin complex as a function of the number of moles of platinum added per mole of 1.02×10^{-5} M protein at pH 7.77.

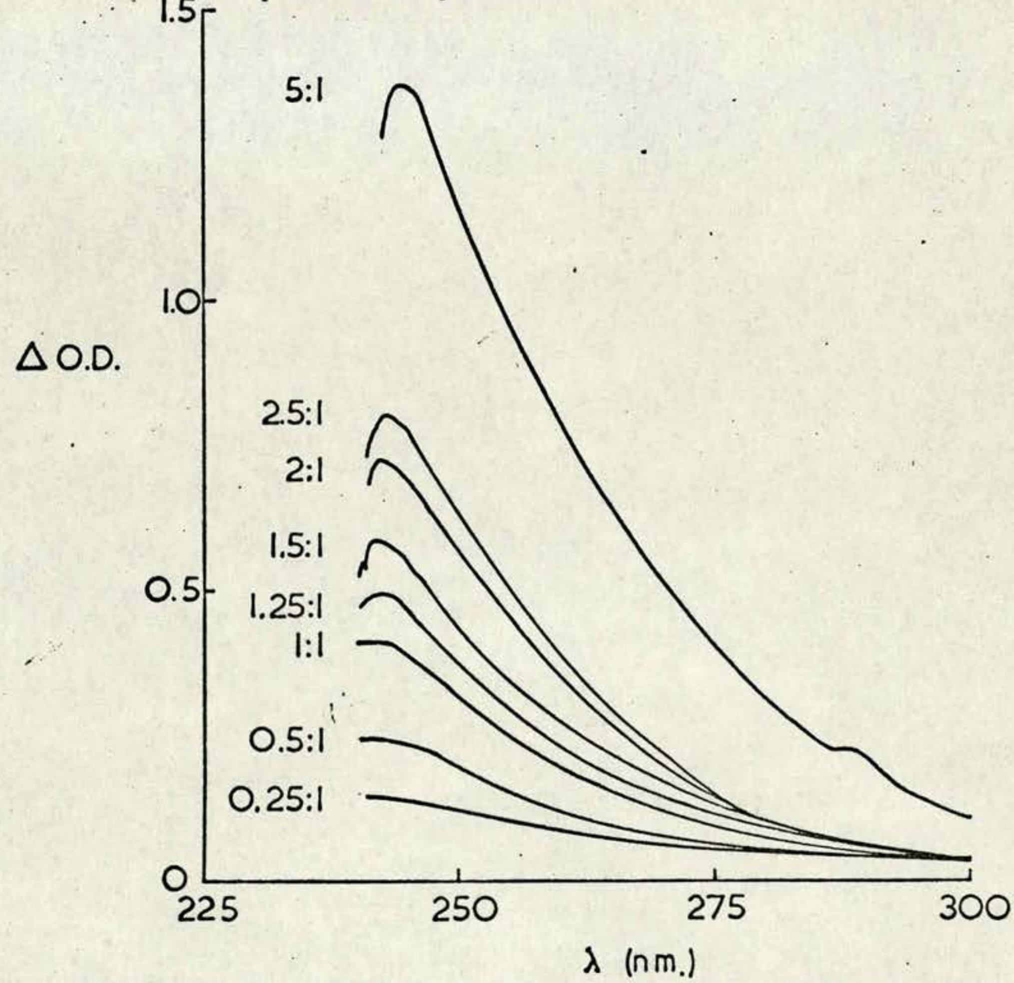


FIG. VIII.6b Difference of the solutions from Fig. VIII.6a recorded on the SP800. The reference was the same concentration of protein.

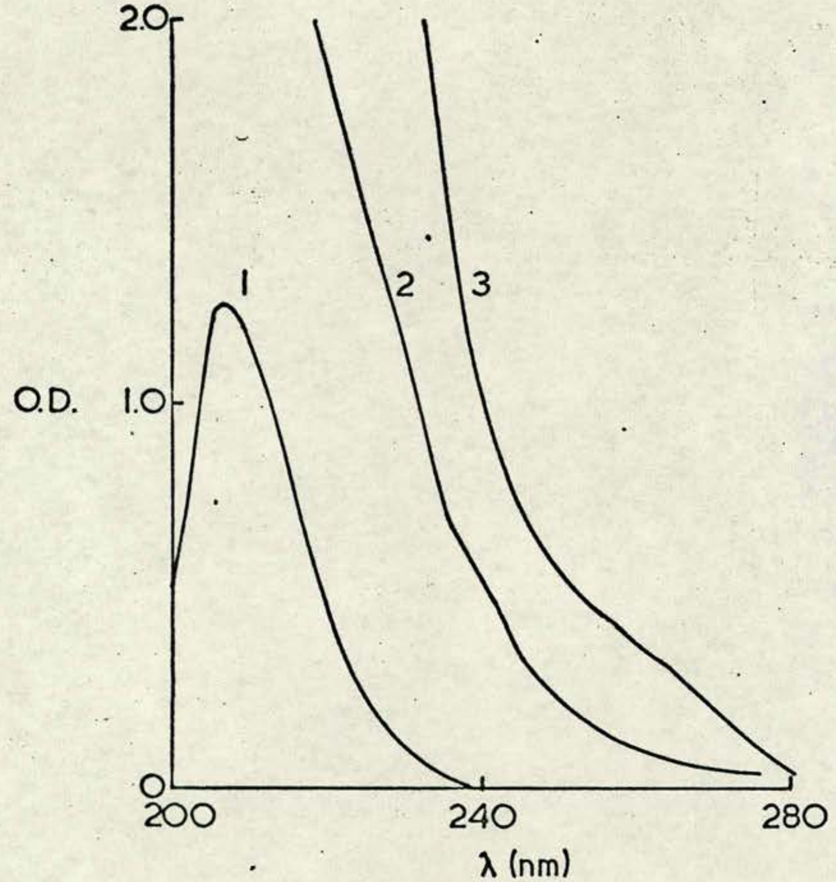
it could be that an intermediate stage between no change and too great a change is only attainable after very long periods of soaking the crystals in small quantities of the platinum complex.

In lattice Z, the results were similar. At pH 7.8 and using both ammonium and mixed sodium-potassium phosphates as supernatants at 2.2M concentration it was found that very little change occurred after three weeks' soaking in a 10mM solution of $PtCl_6^=$. However, after three months, the crystals had lost most of their birefringence and the diffraction pattern did not extend beyond about 4Å. The cell dimensions also had changed to those shown below.

	a(Å)	c(Å)
Native	54.4 ± 0.3	113.1 ± 0.7
Derivative	54.6 ± 0.4	114.9 ± 1.0

With tetrachloroplatinite (II), whose spectrum is shown in Fig. VIII. 5 the experiments above were repeated and very similar results were obtained. Once again, the reaction progresses further than it does with TCA, as can be seen from Fig. VIII. 6a and b. There was also the gradual increase of the specific rotation with time noticed with the $PtCl_6^=$ solutions.

Both combined equilibrium constants for the tetrahydroxy- and tetraammine-platinite complexes are greater than twice that of the $PtCl_4^=$ ion. It is probable, therefore, that the observed reactions were not only with $PtCl_4^=$ but also with mixed complexes of Cl^-, NH_3 and OH^- (the pH being 7.8). The NH_3 would be most likely to alter the binding characteristics because the charge would change. It must also be remembered that Dickerson et al.



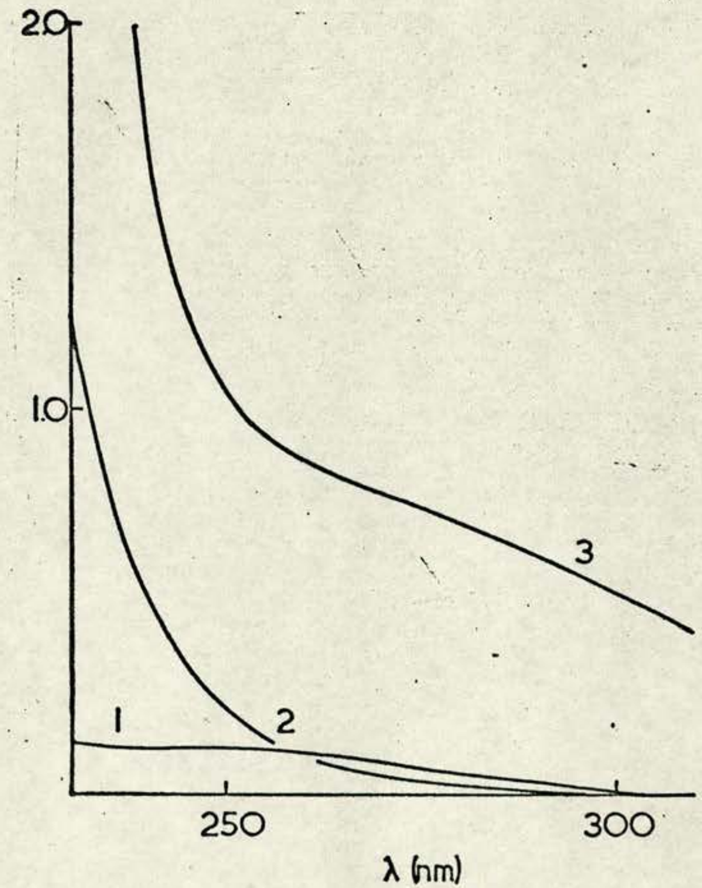


FIG. VIII.8 As for Fig. VIII.7 but with a 1.0×10^{-5}M methionine solution replacing that of cysteine.

FIG. VIII.7 The reaction of $PtCl_4^=$ with cysteine at pH 7.8. The cysteine was 1.0×10^{-5}M and the Pt was present as a 5:1 excess. 1: cysteine alone; 2: $PtCl_4^=$ alone; 3: the complex.

(1969) have found crystallographic evidence for the binding of $PtCl_4^=$ to four kinds of site. All of this makes the results of an examination in solution of the platinum complexes interaction with proteins rather difficult to interpret unless the residues involved are actually known. This requires a knowledge of the structure.

In order to see if any reaction could be detected by u/v spectroscopy, preliminary experiments were carried out with a fivefold excess of platinum over, in the first case, cysteine and, in the second, methionine. The spectra obtained are shown in Figs. VIII. 7 and 8 and it will be seen that definite changes do occur. These require further investigation into their stoichiometry.

The X-ray evidence with β-lactoglobulin lattice Y crystals is, as in the case of $PtCl_6^=$, that at low concentrations of heavy atoms no changes seem to occur whilst, at higher ratios (3:1 metal to protein and above) large changes, fall-off and distortion result. The cell dimensions change to

$$a = 55.8 \text{ Å} ; \quad b = 67.7 \text{ Å} ; \quad c = 82.9 \text{ Å}$$

which are almost identical with those of the $PtCl_6^=$-lactoglobulin derivative. Furthermore, the two (Okl) diffraction spectra are very similar indicating that the binding sites are very close together if not actually coincident. A similar situation has been reported by Sigler and Blow (1965) in α-chymotrypsin where $PtCl_6^=$ in phosphate buffer occupies exactly the same sites as $PtCl_4^=$ in A.S.

The lattice Z work was disappointing in that only very small changes were obtained in the 10mM diffusion preparation which was photographed some three weeks after it had been set up. This, however, was purely preliminary and trials should not be abandoned since a fresh preparation using stoichiometric amounts of the heavy atom could well give good results if left for a longer period to attain equilibrium.

The chloro complexes of platinum, then, pose several questions in connection with their use as heavy atom compounds for isomorphous replacement studies. For example, much solution work on $PtCl_4^{=}$ has been reported in the literature (see the review by Basolo and Pearson (1967)) but the ligands reported have been uncomplicated by protein standards. There might be some benefit, therefore, in examining the proteins whose binding sites have been determined crystallographically, in an attempt to tie together their solution behaviour and the known side-chain interactions.

Iridium Complexes.

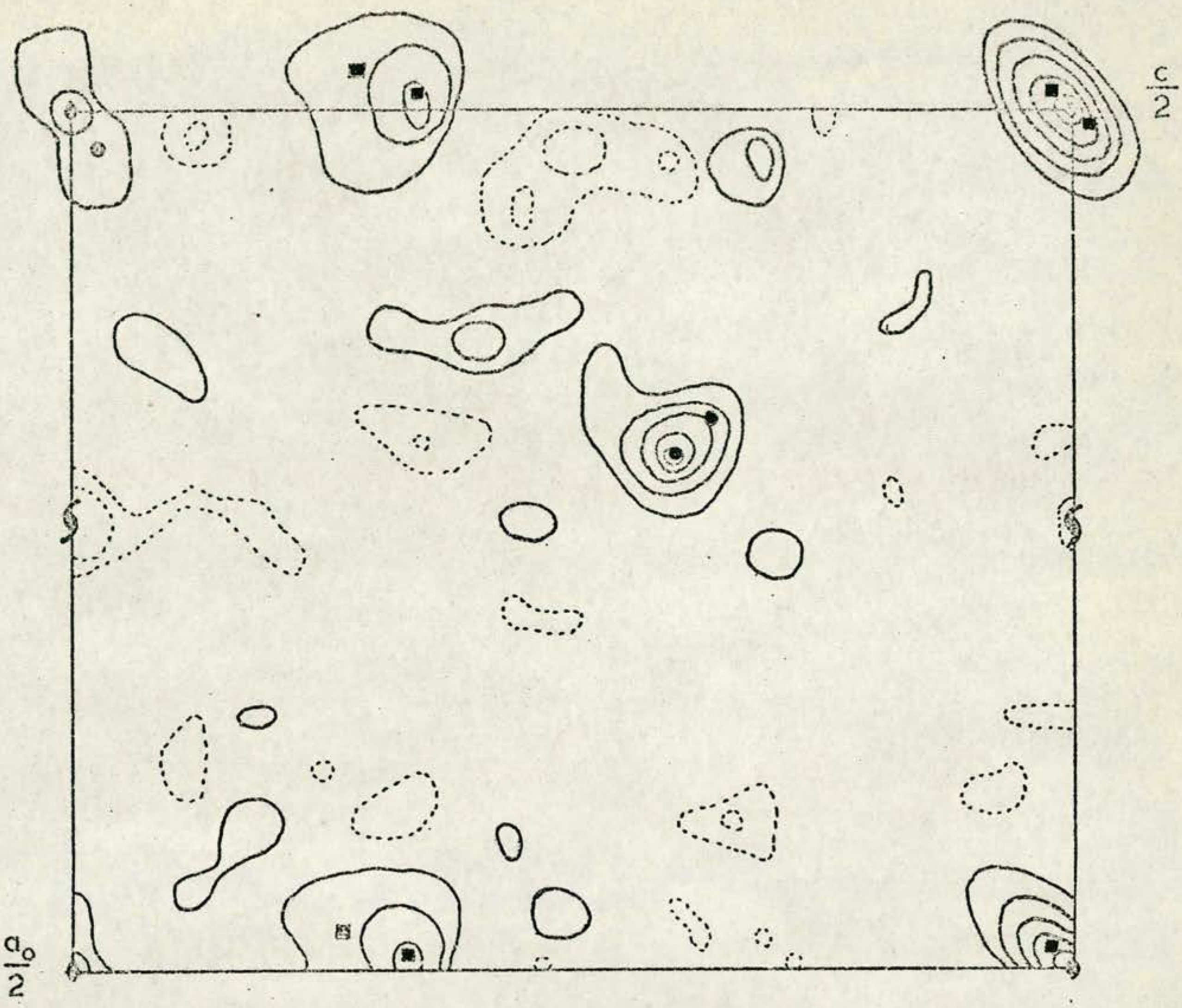
Isomorphous with $(NH_4)_2PtCl_6$ is $(NH_4)_2IrCl_6$ (Wyckoff (1965)). Both complexes are in an upper oxidation state but, whereas the platinum one changes to a square-palnar configuration on reduction, the iridium one remains octahedral. The rates of acid hydrolysis of the $IrCl_6^{3-}$ species have been determined by El-Awady, Bounsall and Garner (1967) as 9.4×10^{-6} sec.$^{-1}$, 8.0×10^{-7} sec.$^{-1}$ and 8.0×10^{-9} sec.$^{-1}$ for the addition of the first three successive water molecules. An equilibrium constant for the first step has been

$$IrCl_6^{3-} + H_2O = IrCl_5(H_2O)^{=} + Cl^{-}$$

determined in the same laboratory as 5. For the Ir(IV) species no equilibrium or kinetic data are available although Garner and his coworkers show spectra of the trichlorotriaquo complex so that hydrolysis obviously does occur to an appreciable extent.

Using the same molarities of β-lactoglobulin and both $IrCl_6^{=}$ and $IrCl_6^{3-}$ as had been used for the platinum solution studies, no changes were observed to occur either in u/v spectrum or in specific rotation. This latter result showed that no binding which disrupted the conformation was occurring. This, in turn, was taken to indicate that no reaction was occurring at the sulphydryl site. A fairly delicate indication of substitution of the chlorides attached to the iridium atom might be obtained from an examination of the visible and near u/v spectra, judging from the work of Garner and his colleagues, but the peaks in question have very small extinction coefficients which would necessitate the use of high concentrations of protein and complex. Whilst this was not impossible, it was uncertain whether the protein bound the iridium complexes at all so that it was decided to press on with the diffusion trials before returning to the solution studies. The lack of any change in the specific rotation also pointed to the possible substitution at sites other than the MMA one.

A 9^{o} [010] precession photograph of a lattice Z diffusion trial with $IrCl_6^{=}$ in sulphate had been taken by Green. This showed large changes affecting the spectrum to the limit of the photograph (about $\underline{6}$Å resolution). The data obtained from this



<u>FIG. VIII.9</u> Orthogonal [010] lattice Z ΔF Fourier projection for $IrCl_6^{=}$ to 6Å resolution. The phases are those calculated from the MMA and PTN derivatives and the three-fold screw axis lies along c. The related sites are marked and the zero contour has been omitted. Contours are drawn at 2.0 electrons/Å^2.

	Initial	16 Cycles	10 Cycles	9 Cycles	9 Cycles	10 Cycles	12 Cycles
Scale	1.000	-	-	0.887	-	-	0.877
Occupancy	80	80	-	-	80	-	98
B	20.0	-	54.13	-	-	45.88	46.75
x	0.102	0.113	-	-	0.113	-	0.115
y	-0.346	-0.364	-	-	-0.364	-	-0.365
z	-0.002	-0.009	-	-	-0.006	-	-0.002
Occupancy	48	49	-	-	56	-	76
B	20.00	-	19.78	-	-	70.25	70.69
x	0.984	0.984	-	-	0.988	-	0.999
y	0.822	0.857	-	-	0.859	-	0.870
z	0.156	0.156	-	-	0.156	-	0.157
R.M.S..Del	210	201	201	184	185	186	177
R	0.654	0.646	0.649	0.622	0.633	0.630	0.567

219 Reflections.

TABLE VIII.1 Refinement of the two sites in lattice Z for the $IrCl_6^{=}$ derivative.

photograph were used to calculate a difference Fourier projection with the phases from the two derivatives already measured (Simmons (1965)). This showed there to be occupation at, or at least close to, the two HGI sites. In case this was a spurious result obtained by use of the HGI phases, a fresh difference Fourier map was calculated based on the phases from the MMA and PTN derivatives. This is shown in Fig. VIII. 9. Table VIII. 1 gives the stages of refinement by PANGLOSS and the final sites were found to be about 5Å and 2.2Å distant from the HGI major and minor sites respectively, the major change in both sites being in the y-coordinate which is in the direction of the dyad. The occupancy can be seen to be above 1.0 indicating that electrons over and above the full occupancy for iridium are present. Also, the temperature factors for both sites are large indicating a "smearing" of the added electrons consistent with the addition of a complex rather than a single atom. The cell dimensions were measured as

$$a = 54.7 \pm 0.4 \text{ Å} \quad ; \quad c = 112.7 \pm 0.5 \text{ Å}$$

and therefore this derivative looked as though it might be a suitable one to replace that of HgI_4^{2-}, provided that there was no great fall-off at resolutions better than 6Å. Unfortunately, however, no note of the preparation could be found so that it was not known how much of the iridium complex had been added per protein molecule. Diffusion trials were therefore repeated.

Crystals of lattice Z were soaked in 3mM, 10mM and 20mM $Na_2 IrCl_6.6H_2O$ and 2.4M potassium phosphate buffer at pH 7.8. After a month the most concentrated sample showed only very small

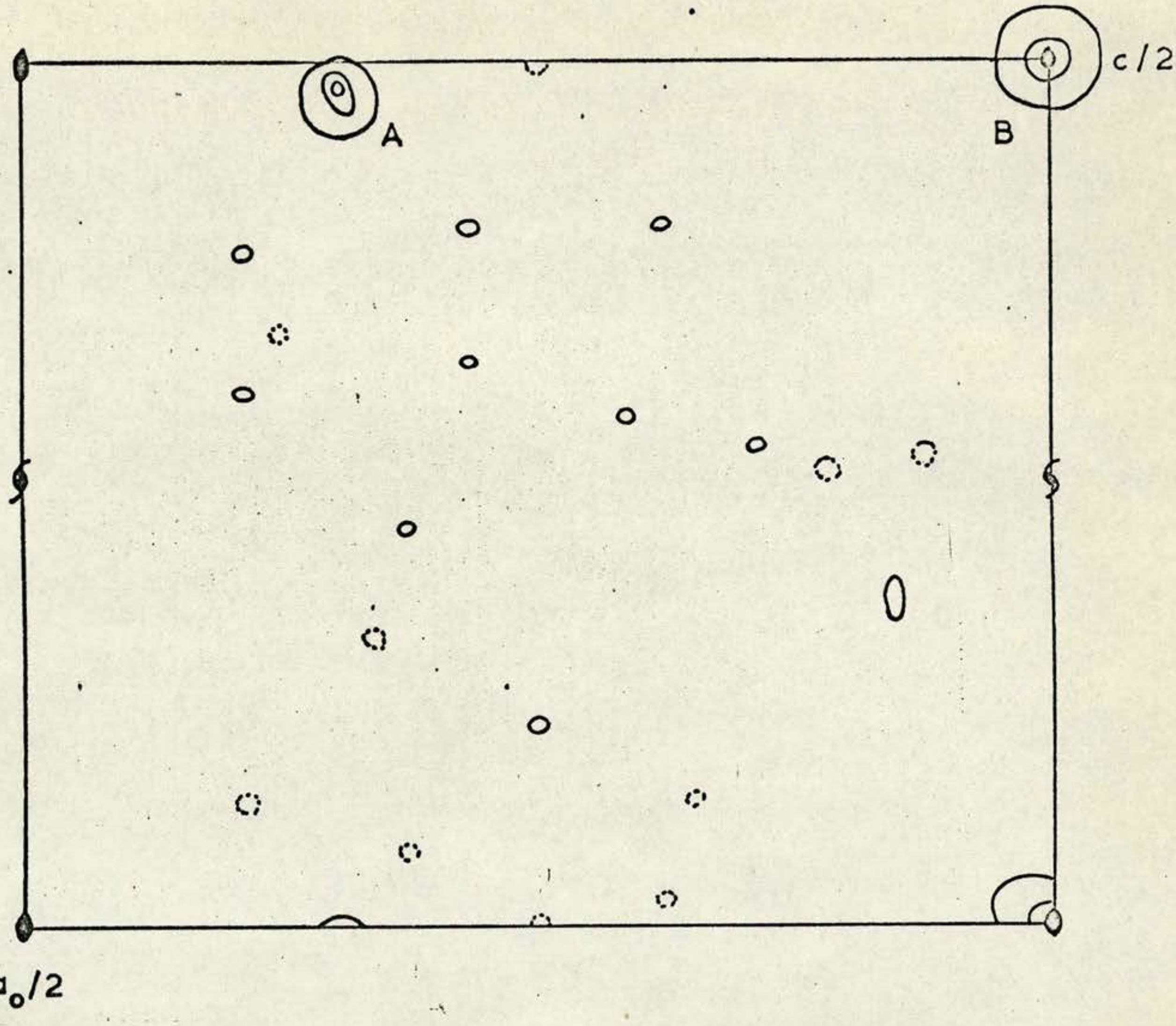


FIG. VIII.10 The [010] difference Fourier projection of the $IrCl_6^{3-}$ derivative using 3.5Å data and phases from the MMA derivative. Contours are at intervals of 2.0 electrons/Å^2.

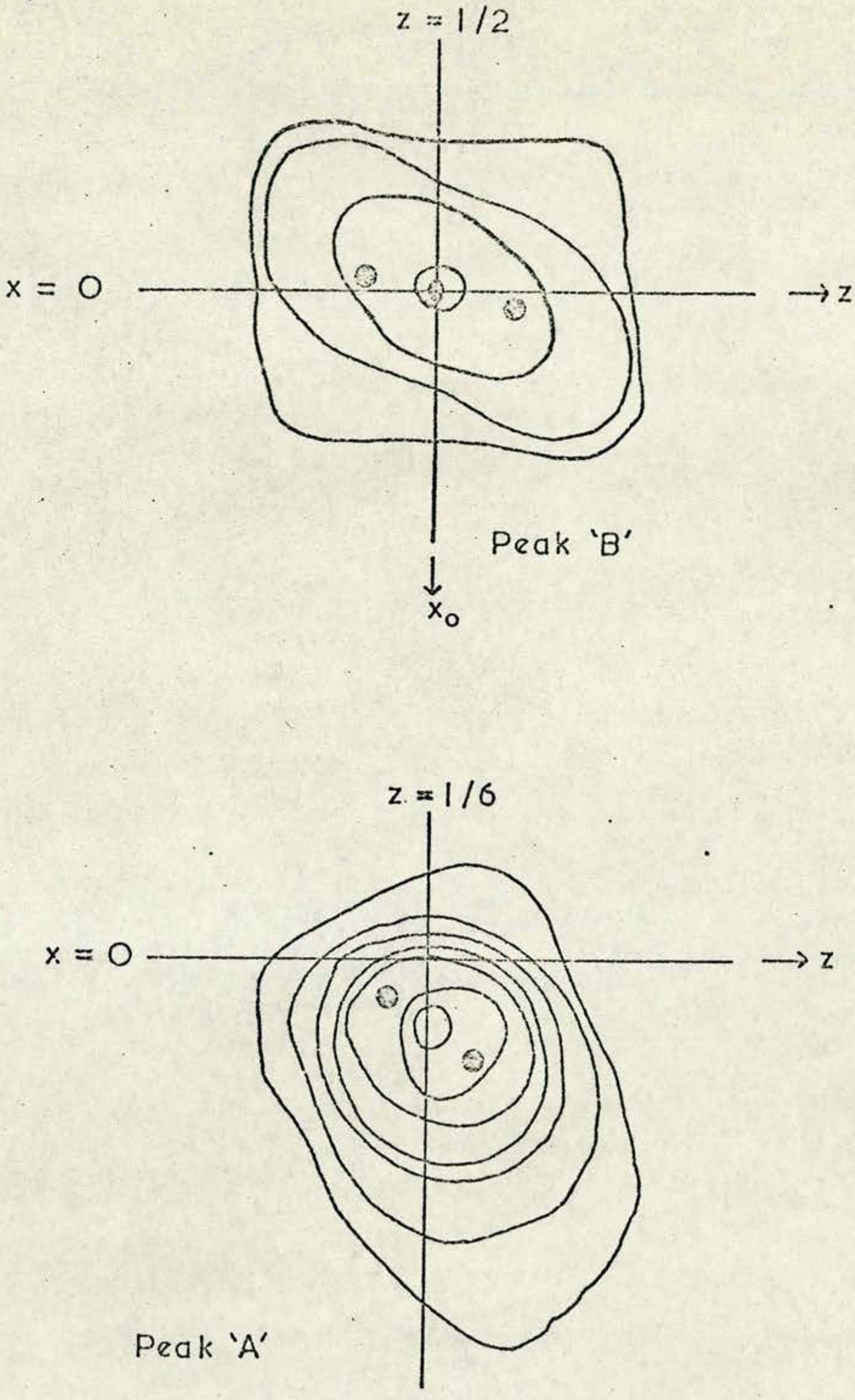


FIG. VIII.11 Enlargements of the peaks A and B from Fig. VIII.10 with contours at 2.0 $e/\AA^{3}$.

	Initial	9 Cycles	14 Cycles	19 Cycles
Scale	1.000	-	0.966	-
Occupancy	40	31	-	32
B	20.0	-	-	14.0
x	0.01667	0.03667	-	0.03667
y	0.00833	0.01708	-	0.01608
z	0.16667	0.16473	-	0.16485
R.M.S, Del	73	63	62	62
R	0.692	0.610	0.609	0.605

560 Reflections.

TABLE VIII.2 PANGLOSS refinement of the $IrCl_6^{3-}$ data to 3Å.

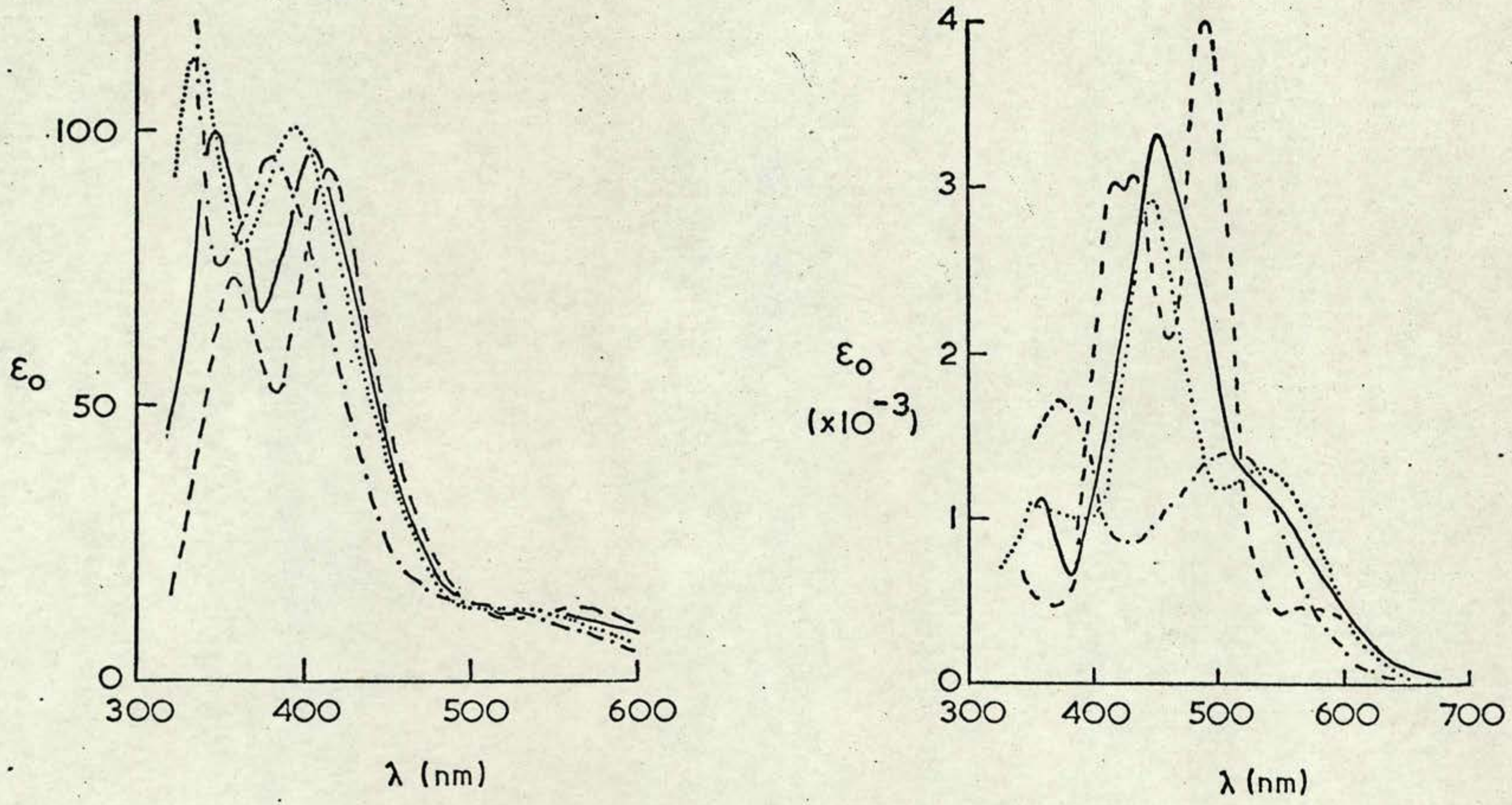
changes and these were at low resolution. However, $Na_3IrCl_6.12H_2O$ under the same conditions save that ammonium phosphate buffer was used, gave small but significant changes all the way out to the limit of a 13^o precession photograph (about 3Å). This was with the 20mM diffusion trial.

Consequently, the [010] photograph showing these changes was measured using the microdensitometer perpendicular to the c-axis in order to minimise the error in the base line from which the peak heights were measured. The raw intensities were corrected for Lorenz and polarisation effects and then a difference Fourier map was produced by means of COLLATOR using the phases calculated from the MMA derivative, this being the only one for which 3Å data were available. Some twenty or so low order terms were omitted because the changes of these reflections were much larger than the average and arose from the change in salt medium in the tube in which the diffusion took place.

The difference Fourier map so calculated is shown in Fig. VIII. 10. The sites were found to correspond approximately to the HGI minor site and a $(\Delta F)^2$ synthesis showed peaks which were consistent with this being the only site. Because of the close proximity of the two peaks in projection, enlarged sections of these areas were calculated (Fig. VIII. 11). This enabled parameters to be obtained for the refinement of the site with PANGLOSS. Table VIII. 2 lists the stages of refinement giving also the final parameters. As can be seen, the final occupancy was low with only 55% of an iridium atom present but what was more noticeable is that the site was some 9.5Å distant from the HGI minor site, once again most change occurring in the direction of the molecular dyad. In this case, however, the difference was more than twice the complex's radius which Wyckoff (1965) gives as 3.4Å. This meant that the binding could not be occurring at the same point as the HgI_4^{2-} ion. Also, no evidence of substitution at or near to the HGI major site was obtained so that the difference in charge between this and the preparation by Green was sufficient to prevent substitution at both HGI sites. In other words, it looked rather is if a new derivative had been formed. Could the occupancy be increased?

Photographs taken some three months later showed a large amount of fall-off, the pattern barely extending to 3.5Å. This indicated that further extensive diffusion had occurred since the first data were collected. It is therefore well worth repeating these experiments with fresh lattice Z crystals and stoichiometric amounts of $IrCl_6^{3-}$ and to allow at least eight weeks for equilibration. Similarly, $IrCl_6^{2-}$ has promise as a replacement for HgI_4^{2-} if the earlier preparation can be repeated.

In addition, it was noted that in both sets of diffusion trials the solution had turned a dark blue colour after some eight to ten weeks. This occurred whether the salt was rich in NH_4^+ or K^+ ions. Some slow solvolysis mustoccur and since both solutions seem to end up as the same species as far as can be seen from the colour, some oxidation or reduction must also be happening. The oxidation potentials of the hexachloro iridium system given by Latimer (1952)are those shown below.



<u>FIG. VIII.12</u> Hydrolysis of $IrCl_6^{3-}$ at 25°C in 2.5F $HClO_4$ - 1.2F $NaClO_4$ solution. --- $IrCl_6^{3-}$; —— $Ir(H_2O)Cl_5^{=}$; ... $Ir(H_2O)_2Cl_4^-$; -·- $Ir(H_2O)_3Cl_3$.

Hydrolysis of $IrCl_6^{=}$ at 25°C in Cl_2-saturated 2.5F $HClO_4$ - 1.2F $NaClO_4$ solution. --- $IrCl_6^{=}$; —— $Ir(H_2O)Cl_5^-$; ... $Ir(H_2O)_2Cl_4$; -·- $Ir(H_2O)_3Cl_3^+$.

$$IrCl_6^{=} \xrightarrow{1.017v} IrCl \xrightarrow{0.77v} Ir$$

$$IrCl_6^{=} \xrightarrow{0.836v} Ir$$

The tendency, therefore, is towards the reduction of the Ir(IV). However, Latimer gives two reactions which could cause the oxidation of Ir (III).

$$O_3 + 2H^+ + 2e^- \rightleftharpoons H_2O + O_2 \qquad E = 2.07v$$

$$\tfrac{1}{2}O_2 + 2H^+ + 2e^- \rightleftharpoons H_2O \qquad E = 1.23v$$

The latter, if oxidation were the reason for the colour, would probably be slow but since the pH was above 7.0 the amount of H^+ ion would be rather small. Another possible explanation is that the reduction of the complexes to Ir(O) can give a colloidal suspension which is blue. This might be the more likely explanation judging by the oxidation potentials and the pH of the solution.

That solvolysis occurs to a marked extent with chloro-iridium complexes can be seen from Fig. VIII. 12 where the various spectra of the hydrolysed Ir(III) and Ir(IV) complexes of chloride are illustrated. $Ir(IV)Cl_3(H_2O)_3$ can be seen to have a minimum at about 420nm, the peaks on either side having extinction coefficients some two to three times as large (El-Awady, Bounsall and Garner (1967)). The trichlorotriaquo species is therefore most likely blue. The hydroxide, $Ir(OH)_4$ is also blue but it is insoluble so that an explanation in terms of either solvolysis or complete reduction seems more likely.

Once again, though, there is the problem of the actual species which is binding to the protein and here is visual evidence of a slow change of the species which are present in solution. Could the reason for the high fall-off after ten weeks' diffusion be that the $IrCl_6^{3-}$ is binding to the site determined but, with time, air oxidation (or whatever process is occurring) causes a series of compounds each of which can bind in a separate place because of its different charge, and hence multiple substitution occurs with the subsequent disruption of the lattice?

Camerman and Green (unpublished results) have found evidence of multiple substitution in lattice Y crystals soaked in K_2IrCl_6 at a heavy atom-protein ratio of 6:1. The major site, 60% of an iridium atom was 3.1Å from the HGI site in lattice Y and there appeared in a "double-difference" Fourier map some four other sites with occupancies varying from 12% to 40% of an iridium atom. This diffusion, too, seemed slow to take up the chloro-iridate and possibly a 2:1 ratio left for at least six months might give only a single site.

It remains, then, to draw the various strands together to form an overall picture of the possibilities of linking studies in solution to those in the solid state in an attempt to systematise the finding of suitable heavy atom derivatives with which to produce a high resolution map of β-lactoglobulin in particular and proteins in general. This will be done in the next and final chapter.

CHAPTER IX

This chapter can be roughly divided into two parts: the first deals with the work in the latter part of this thesis and other crystallographic studies on β-lactoglobulin and the second, more general section will deal with some aspects of protein structure determination which have been raised by the work as a whole. It is hoped that in this way the work on the two separate proteins will be brought more closely together.

Discussion of the Results from β-lactoglobulin Studies.

Perhaps the most notable fact which emerges from the work, both in this thesis and by Green and his coworkers, is that there seem to be three major binding sites in β-lactoglobulin, although several other minor ones exist to a greater or lesser extent depending on the heavy atom. The nature of these binding sites is not clear at present save that one appears to be half buried and has the free sulphydryl group associated with it. From the crystallographic point of view, however, the following derivatives look hopeful, in addition to MMA, for high resolution work provided fresh preparations are made under carefully controlled conditions. It appears that lattice Z may be the better of the high pH forms because of the possible exploitation of the HGI minor site. $IrCl_6^=$ looks promising in this respect, as does TCA where the HGI sites are both occupied but with the occupancies reversed. TCP should be tried, also, in lattice Z. $IrCl_6^=$ is another derivative which might well give useful high resolution data. $Pt(CN)_6^=$, if made properly so that it was known to be hexacyanated, $PtCl_4^{=}$ and $PtCl_6^{=}$ are three heavy atoms which should be tried again despite the rather disappointing results obtained. initially. Fresh batches of lattice Z crystals containing small stoichiometric ratios of heavy atom to protein of these compounds should be set up and allowed to equilibrate for a period of several months so that the occupancies can be checked and in the hope that the smaller quantities of additive will not cause non-isomorphism.

This last paragraph raised the point of the occupation of the three main sites so far found: are there common factors about the complexes used which select any one site rather than another? Is the overall charge important? Are the sites just 'general holes' in the surface of the protein or areas between the subunits?

The easiest site to consider is the sulphydryl one where it was shown that direct, covalent reaction occurs. There are two possible reasons why the TCA does react and the similarly-shaped TCP does not, the final answer probably being a combination of the two. First, the overall charge increases from TCA to TCP and this could be responsible for sufficient electrostatic repulsion from neighbouring groups to keep the TCP from binding. Second, the 'driving force' for the reaction with TCA is absent in the case of TCP. It is hoped that the final three-dimensional structure will be able to furnish the answer.

Turning to the HGI site, any speculation as to its nature must of necessity be less precise. Part of the problem is that it is not certain which of the possible HgI_x^{2-x} for x between 2 and 4, is present. The site appears to be close to the molecular dyad so that it is possible the complex is jammed in the crevice between subunits, being bound through interactions with the readily polarisable iodines. This binding causes the non-isomorphism which makes the derivative useless for high resolution work in both the high pH forms although, in lattice X, it forms the most acceptable derivative (Komorowski 1971)). Other compounds which associate at this site include TCA, though surprisingly not TCP, and $IrCl_6^{=}$. These appear to have different charges so that in this case it might be supposed that the charge is unimportant,but $IrCl_6^{=}$ is known to exchange with H_2O and this would reduce the charge by one or more. Might it even be argued that since TCA does not solvate and does bind, and TCP likewise does not solvate nor does it bind, that the site requires a single negative charge obtained by loss of a ligand in both $HgI_4^{=}$ and $IrCl_6^{=}$? However, once more no real conclusion as to whether the charge is an important feature of the site can be obtained without the full structure.

Finally, the PTN site appears to bind square planar complexes of Pt(II) such as $Pt(NO_2)_4^{=}$ and TCP. From this point of view the site seems to be more specific than the previous one although the position is less well defined. This could arise from its being on the outer surface of the molecule and therefore requiring a rather special arrangement of ligands and overall charge for binding to the various polar side chains in the region. As TCA does not bind at this site the charge would appear to be an important factor in this case.

Another interesting point is the way in which the various diffusion trials of lattice Y by Green and his colleagues, and in this work have led to the observation that occupation of sites seems to be an "all-or-nothing" event. Either diffusion occurs rapidly and affects the diffraction spectra markedly, producing not only large intensity changes but also greater fall-off and lattice distortion, or almost no effect is detected. This should be overcome by repeating the trials using intermediate amounts of heavy atom but allowing a long time for equilibrium to be reached. This also applies to lattice Z where similar "all-or-nothing" effects have been observed. In point of fact, lattice Z looks more promising for high resolution study because the plane of heavy atoms is not on or near a special position and also because the substitution of the HGI minor site looks hopeful for a fourth derivative. The production of derivatives in this manner, although slow, also has the advantage that the crystals and mother liquor are in equilibrium so that all crystals required for a data set do not have to be mounted at the same time to reduce the chance of collecting data from unevenly substituted specimens.

The main objection to this approach, and it is only an objection if correlation between the species bound and the binding site is sought, is that the long times involved in the diffusion allow slow solvolysis of the heavy atom to occur. The paper by Sigler and Blow (1965) illustrates this point exactly, in the case of the ammonolysis of $PtCl_4^{=}$ in A.S. solution: except the binding to α-chymotrypsin is inhibited rather than its position altered. When this occurs they suggest a change of mother liquor to remove the source of interfering ions. Alteration of the pH may also serve to remove a source of interfering substituant by altering an equilibrium constant. However, water is the offending molecule in the binding of chloroiridium complexes to the high pH forms of β-lactoglobulin and this makes the problem insuperable until the structure is known. This is because if solvolysis of the complex changes its binding site on the protein, then multiple substitution will occur, probably rendering the derivative useless. The whole problem might well be worth investigating since it is known to be possible to produce substituted crystals of β-lactoglobulin with both $IrCl_6^{3-}$ and $IrCl_6^{=}$, but the methods for such a study seem obscure. Is a large excess of heavy atom for a short time better than a 2 or 3 to 1 ratio for a longer time, or vice versa? If one is better than the other then what exactly are the species involved? This problem may also arise with other complexes of heavy atoms used for the isomorphous replacement method.

What conclusions can be reached from the studies of β-lactoglobulin in solution as an aid to the X-ray crystallography? In this respect, the work with TCA was a good example of a reaction which occurred and could be followed to completion both polarimetrically and spectrophotometrically and in which the X-ray evidence bore out the findings. It was fortunate that the u/v spectrum of TCA was such that it did not hinder absorption measurements. Also, the Tanford transition was of great value in this particular case in that it gave an easily monitored standard curve as a function of pH which served as a yardstick against which to examine further structural alterations. It is interesting that in lattice Z, TCA showed no evidence of reaction at the sulphydryl site and this shows that caution always should be exercised in the extrapolation of solution data to the solid state. The three dimensional structures of lattices Y and Z will show why there is a difference in the mode of association of TCA between the two forms. Before leaving TCA, it is worth mentioning that here is another reagent which may be used to incorporate a heavy atom at a sulphydryl site and it would be well worth trying this complex with other proteins in order to see exactly how specific it is. There are a number of mercurials which are equally suited to cysteine labelling but perhaps the different shape of the TCA molecule might allow reaction to occur where no mercurial reaction occurs. Also, it is unlikely that substitution will occur in exactly the same place so that it might well be used in conjunction with mercurials to solve the phase ambiguity.

Work with TCP was not nearly so convincing because of the absorption band's occurring on top of those of the protein. However, here the specific rotation data showed there to be no reaction at the sulphydryl group and this was borne out by the X-ray evidence. Similarly, other correlation studies were unsuccessful not only because of the difficulty with the overlapping absorption spectra but also because the absorption bands in the visible region were so weak that the changes known to occur on substitution of the metal complex ligands would escape detection at the concentrations employed. The real problem of a correlation between solution and crystal studies, therefore, was and still is that not enough specific heavy atom reagents are known which can attach to groups other than sulphydryl and which have distinctive spectra either before or after.

Some General Considerations.

The need for more specific reagents for reaction to protein side chains is essential if speedy progress in structure determination is to be made. To date, the most successful method of obtaining derivatives is by trial and error, guided to some extent by those compounds which have already been used to advantage in other proteins. This method is slow and requires moderate amounts of protein which may be wasted if the trial causes denaturation. Therefore, if several sidechain specific reagents were available whose reactions could be followed in solution, then these could be being tested on small amounts of material whilst attempts were being made to grow suitable crystals for X-ray work. Thus, when the crystals were available, the trial and error search for derivatives would be minimised and the amount of protein used small. Also, when the interpretation of the electron density map has to be done, certain "markers" would be available by which fitting of the sequence to the contours would be facilitated.

It is possible, of course, that some of the heavy atom reagents in current use <u>are</u> binding specifically and it is just the detection of this which is at fault. The recent discovery by Dickerson et al. (1969) of the binding of $PtCl_4^{=}$ to methionine is an example of a specific reaction which might be used although other reactions also occur so that a trial protein would have to be chosen with care. Perhaps ORD and CD measurements might be able to show up associations, especially the more "internal" ones, which specific rotation data are not precise enough to detect. Nuclear magnetic resonance is another tool with which to investigate the interactions of protein and heavy atom. Here the replacement of a proton by a ligand or even a change in hydrogen bonding might be detected if not by the direct disappearance or shift of a peak then by a change in the chemical shift and splitting of neighbouring protons. I/r spectroscopy, too, could well show up binding as a change in the frequency of certain bands associated with either the protein or the heavy atom complex (or indeed, both). Studies with enzymes have the added advantage that they are known to bind a substrate. If this can be modified and "frozen" into the active site then this affords an obvious way of attaching heavy atoms to the protein (for example, the work of Blow and his coworkers on α-chymotrypsin). However,as more structures are solved and the binding of the various heavy atoms is determined directly from the tertiary structure, it may become clearer which heavy atoms are specific and which are not and it may even become possible to determine which methods would be most likely to be of use in a solution study so that the protein-ligand binding can be extended to other proteins. As an example, ligand X has been used to solve the structure of a protein, P, and, from the structure has been found to associate with a particular residue (or group of residues). In solution, a method of detecting such an interaction is sought so that, if a positive result is obtained with a second protein, E, under the same conditions, then there is a good chance that X will be found to associate with E in a similar manner as in P, assuming that no blocking occurs because of the lattice.

In chapter III, the factors affecting the growth of crystals for X-ray study were discussed. Finally, therefore, some general principles on how to approach the problem of producing good crystals will be discussed since this has become the rate-limiting step in tertiary structure determination because, without them, there is, as yet, no other satisfactory method. Ideally what is required is a crystal form which is stable at room temperature, which is quickly and reproducibly prepared and which has a long life in ionising radiation so that as much data as possible can be collected from the one crystal. The derivatives should be easy to prepare, isomorphous and have only one or two sites of substitution whose occupancies are the same from crystal to crystal. The unit cell should be relatively small and should contain a small asymmetric unit. Finally, the primary sequence should be known. Unfortunately, all too often few, if any, of the above requirements are fulfilled.

What factors are likely to be of importance in the production of large crystals? Reviews by Czok and Bücher (1960) and Dixon and Webb (1961) discuss the effects and relative importance of the parameters affecting crystallisation: as mentioned before, these are the concentrations of protein and salt, pH, temperature, the nature of the salt and, if present, the organic solvent. Blake (1968) discusses the preparation of isomorphous derivatives and the methods for doing this which are currently available. However, the process of producing crystals of a protein which are suitable for a full X-ray investigation still is empirical to a great extent.

Perhaps the most awkward point which arises at the start of an attempted crystallisation is the doubt there is concerning the purity of the protein. What complicates matters is that any impurity may well have very similar solubility properties to the required protein, especially if only one procedure has been used in the purification or, more seriously, if some minor modification such as the hydrolysis of an asparagine or glutamine has occurred during fractionation. Thus, with proteins, crystallinity is no real criterion of purity and should be regarded with suspicion especially if small (the "silky sheen") or poorly formed crystals occur. Therefore, other yardsticks should be used in conjunction, such as electrophoresis, the specific activity (if applicable) and ultracentrifugation. If these criteria are all met then the protein may be judged "pure". An extra point which is well worth noting is that brought up by Dixon and Webb (1964). Certain lipids seem to be able to escape entire removal by many of the separation methods which are normally used and therefore an organic solvent precipitation is often beneficial in removing such contaminants and allowing crystallisation to occur where before it would not.

The solutions used may contain metallic impurity which will bind to the protein and may well interfere with the crystallisation. The addition of small quantities of EDTA during recrystallisation of the salts used for making up the various solutions has been suggested by Czok and Bücher as a precaution. The recrystallisation of these salts is vital since even analytical grade commercial products contain relatively large amounts of insoluble material which are likely to interfere with the crystallisation by supplying many unwanted nuclei. Clean glassware and a supply of glass-distilled water, preferably filtered through a very fine filter (Millipore) are also essential. Finally, the method of addition of the protein is important since, if run down the sides of the tube, the crystals often form attached to the walls from which they are difficult to remove intact. The protein is usually added dropwise onto the surface of the salt solution.

These, then, are the precautions which should be taken in order to ensure that contamination is kept to a minimum. How can the next stage best be carried out: that of producing large, well-formed crystals? The reasons for preferring the salting-out method were given in Chapter III and it is to this method that the following should be applied. The stratagem described is not intended to be either comprehensive or fool-proof but most of the salient points are presented, as they appear to the author.

First, the literature should be checked in order to find out if crystals of the protein under study have been obtained before and by what means. This search should also include derivatives of the protein as well as species variants. The information obtained is of value in providing a starting point for the study and may even help in the choice of derivatives. With aldolase, in this work, it was known that there were at least two crystal forms which could be studied. Solubility curves as functions of pH, temperature, concentration of both salt and protein and fraction of organic solvent may also be obtained. These will allow work to be carried out under conditions where the control of the variables is not so vital and where the amount of protein required is a minimum, a point worth considering when preparation is complicated and yields are small.

If no information such as that mentioned above is available then it should be obtained, provided that sufficient material is available. If there is not enough to allow a full series of experiments then possibly the most useful experiment would be the solubility as a function of pH.

Once such information as is readily available has been obtained, then the process of trial and error crystallisation can be started using the pH and concentration values found from the solubility data as an aid to limiting the number of trials. Unfortunately, little more can be done to cut down the number of trial crystallisations which may be required before success is achieved and consequently the process is apt to be lengthy. Holmes and Blow (1966) list the basic principles for obtaining large crystals as

(a) having the minimum number of nuclei present at the start,

(b) only having a slightly supersaturated solution and

(c) maintaining the conditions constant.

The first of these is necessary to stop showers of tiny crystals appearing which will then grow no further. High speed centrifugation of the solution of protein at the point of first precipitation (incipient turbidity) can be used to limit the number of nuclei on which large crystals can grow. The second requirement is to allow sufficient transport between the solid and dissolved states for crystallisation to occur at a very slow rate. Finally, the conditions must be maintained as constant as is possible in order to obtain even growth. Of these, the most important one is the temperature since the concentrations are fixed and the pH is held by adequate buffering. Also, disturbance of the tubes should be avoided during the initial stages since vibration can cause breakage of small crystals thus providing many more nuclei and preventing growth of a few large crystals. Holmes and Blow also state that traces of some impurity, usually in the form of an organic solvent, may be critical in the production of crystals. They cite the requirement of traces of toluene in the production of tetragonal human oxyhaemoglobin crystals. This effect has been observed also by Moews and Bunn (1970) in the case of a rennin-like enzyme from the fungus *Endothia parasitica*.

Of course, the first crystals, once obtained, can be used as seeds for subsequent preparations. Until these are obtained, however, systematic trials varying only one parameter at a time must be set up. If material is in short supply, the process is lengthy. Each protein has different characteristics so that a truly "general" method is impossible although the approach, dictated as it is by the requirements of the crystallographer, *is* gradually becoming more general. As more structural data, both primary and tertiary, become available, it may become more obvious as to which particular starting point is likely to be the most successful with any particular protein, knowing the sequence of the unknown and the structure and crystallisability of related proteins.

Acknowledgements

I would like to thank my supervisor, Dr. D.W.Green for his guidance and encouragement throughout this work. I am also indebted to him for allowing me to quote some of his unpublished results. I am grateful to many members of the Departments of Biochemistry, Chemistry and Natural Philosophy for helpful discussions and for permission to use their instruments and computer programs. Especially I would like to thank Drs. M.M. Harding, D.A. Rees and J.S.McK.McKee and Mr. E.S.Komorowski.

The technical assistance of the workshop staff and, more particularly, Mr. D. Pettigrew is gratefully acknowledged.

Also, I would like to thank Professor N.Feather, F.R.S., for providing laboratory facilities in the Natural Philosophy Department, to the Edinburgh Regional Computing Centre for computing facilities and to the Medical and Science Research Councils for financial support.

Thanks are due to Mrs. Myra Sanderson and Mr. W.Robb for help in the reproduction of the manuscript and diagrams.

Finally, I would like to thank my wife for her patience and encouragement during the writing of this thesis.

APPENDIX I

Refinement of Heavy Atom Parameters by the Method of Hart.

In order to obtain the most accurate phase information from the method of multiple isomorphous replacement, it is necessary to have the most accurate heavy atom parameters. This requires some form of refinement of the values obtained from the $(\Delta F)^2$ syntheses. Because native phase information is not generally available at the start of a protein structure problem, Fourier refinement methods cannot be used. The only phase information available, in fact, is that calculated from the heavy atom site whose parameters are to be refined. This, for non-centric reflections, gives an ambiguity (see Chapter I) in the native phase, resolved usually by the inclusion of at least a second derivative. A least squares procedure then becomes possible, alternating cycles of refinement with those of fresh phase determination (Dickerson, Weinzierl and Palmer (1968)).

In a normal least squares procedure, the function minimised is of the form:

$$R = \sum_{\text{all data}} w.\ (F^{obs} - F^{calc})^2$$

When applied to heavy atom refinement in proteins, the F_H^{calc} term can be found completely by calculation. It is the F_H^{obs} term which provides the problem since only the amplitudes of the native and derivative structure factors are known. Without phases the F_H^{obs} term cannot then be found since

$$\underline{F}_H^{obs} = \underline{F}_{PH} - \underline{F}_P$$

Trial and error methods are clearly out of the question in general but in the case of centric reflections where the phases are either 0 or π , a trial and error method becomes possible. There are four possibilities, then, for the value of $\underline{F}_H^{obs}$ given by

$$+ |F_{PH}| + |F_P| \quad , \quad + |F_{PH}| - |F_P|$$
$$- |F_{PH}| + |F_P| \quad , \quad - |F_{PH}| - |F_P|$$

Those signs are then chosen which give the smallest difference between $\underline{F}_H^{obs}$ and $\underline{F}_H^{calc}$. This is the method described by Hart (1961) and it has been successfully applied to proteins with centric data since then. Lundberg (1965) has described a program which is essentially this save that a full matrix least squares solution is obtained rather than the "one at a time" procedure adopted by Hart.

If there are p sites and n reflections, then the function which is minimised is

$$E = \sum_n w_n \cdot (\sum_p (o_p \cdot f_{p,n} \cdot \exp(-B_p \cdot \sin^2\Theta/\lambda^2) . \cos 2\pi . (hx_p + lz_p) \pm |F_P| \pm |F_{PH}| . k)^2$$

for the h0l projection of space group $P2_1$. The symbols represent the following:

w_n = weight of the nth observation

o_p = occupancy of the pth site of the derivative

$f_{p,n}$ = scattering factor of the pth site of the derivative which is also a function of the angle, Θ, obtained from the hkl for each reflection

B_p = temperature factor of the pth site of the derivative (normally not refined)

x_p, z_p = fractional coordinates of the heavy atom site, p

k = scale factor applied to $|F_{PH}|$

If the method of Lundberg or, indeed, any least squares procedure, is employed the starting parameters must be fairly accurate before convergence is obtained. As this may not be the case with parameters derived from low resolution $(\Delta F)^2$ maps, Hart's method of applying

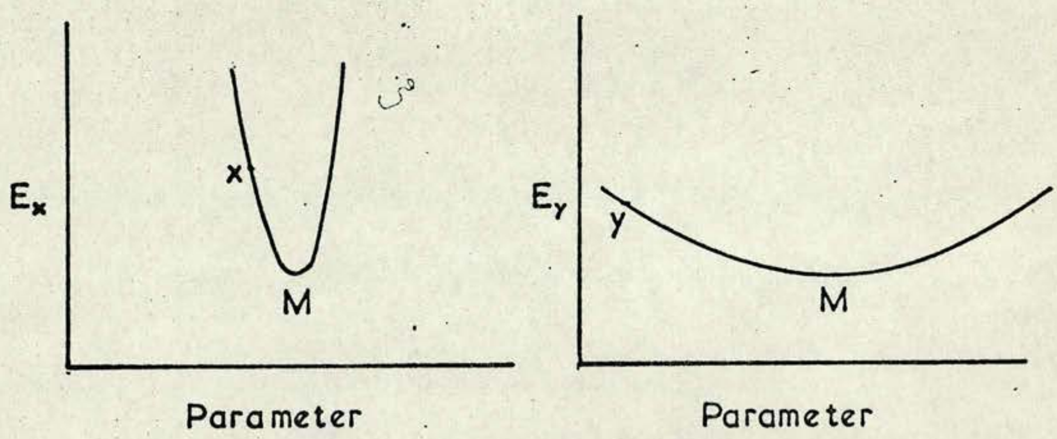


FIG. AI.1 Diagram showing the inadvisability of weighting the shifts for the Hart method of refinement.

fixed shifts, one at a time, to the various parameters is generally more acceptable in the initial stages of refinement.

Hart's method deals with each parameter in turn and applies to it a specified shift -2, -1, 0, 1, 2 times, selecting as the best value of that parameter, that which gives the minimum value of

$$E = \sum (F_H^{obs} - F_H^{calc})^2$$

the summation being over all reflections. This value is then used in the next cycle, the specified value for the shift being altered suitably (see below). At first, Hart applied a weighting scheme such that the parameter causing the largest change in E was given the full shift for the next cycle, others being weighted down according to their effect relative to the largest. This, in fact, slowed down the convergence. From the point X (Fig. AI. I) a small shift will cause a large change in E_x but the value of the parameter is more nearly correct than in the other case where the minimum is shallow and Y is further away from its true position. The minimum, M, is that corresponding to the true values of all parameters and E is the error associated with each parameter alone. Hence, a change of parameter causing a large change in E may not be the most instrumental in completing the refinement and a weighting scheme dependent on the size of E alone will probably slow down the convergence. Hart therefore abandoned the scheme and no such weighting was applied by the author in the present work.

This method of refinement has been described for the more usual crystallographic problems where full-matrix least squares methods have been found to diverge. Four successful applications are given by Bhiuya and Stanley (1963) and PANGLOSS was adapted by the author to refine $(NH_4)HSO_4$ with great success.

PANGLOSS

"All is for the best - -" Voltaire, 'Candide' ch.30.

The program for the work described in the main body of this thesis was written in Atlas Autocode (AA) for the Leo-Marconi KDF 9 run by the Edinburgh Regional Computing Centre (ERCC). This machine had a 48 bit word length and some 16,000 store locations available for both program and data. Peripheral storage was available on direct access magnetic tape and this could cause delays if the sections required were far along the tape. Input was from cards or paper tape and the data tapes produced by the diffractometer which were five-hole rather than the more common eight-hole ones on which most input was made. A suitable reading/translating routine for the five hole tape was available too.

Since the beginning of 1970, however, the ERCC has changed to an IBM 360/50 which, although having a modified Atlas Autocode (IMP) compiler, has a word length of 32 bits, split into 4 x 8 bit "bytes". There are some 50,000 32 bit words available to the user. Back-up storage is provided on direct access discs and also on sequential access magnetic tapes. Input is by cards only although paper tape can be handled, if absolutely necessary, by the ICL System 4-70, at present undergoing acceptance trials at the Centre. An alternative method of feeding data in on paper tape has just become available. A direct link from the Natural Philosophy Department to the 360/50 has been provided through the PDP-8 computer in the former department. This enables the diffractometer data tapes to be processed directly without the tedious copying procedure hitherto required.

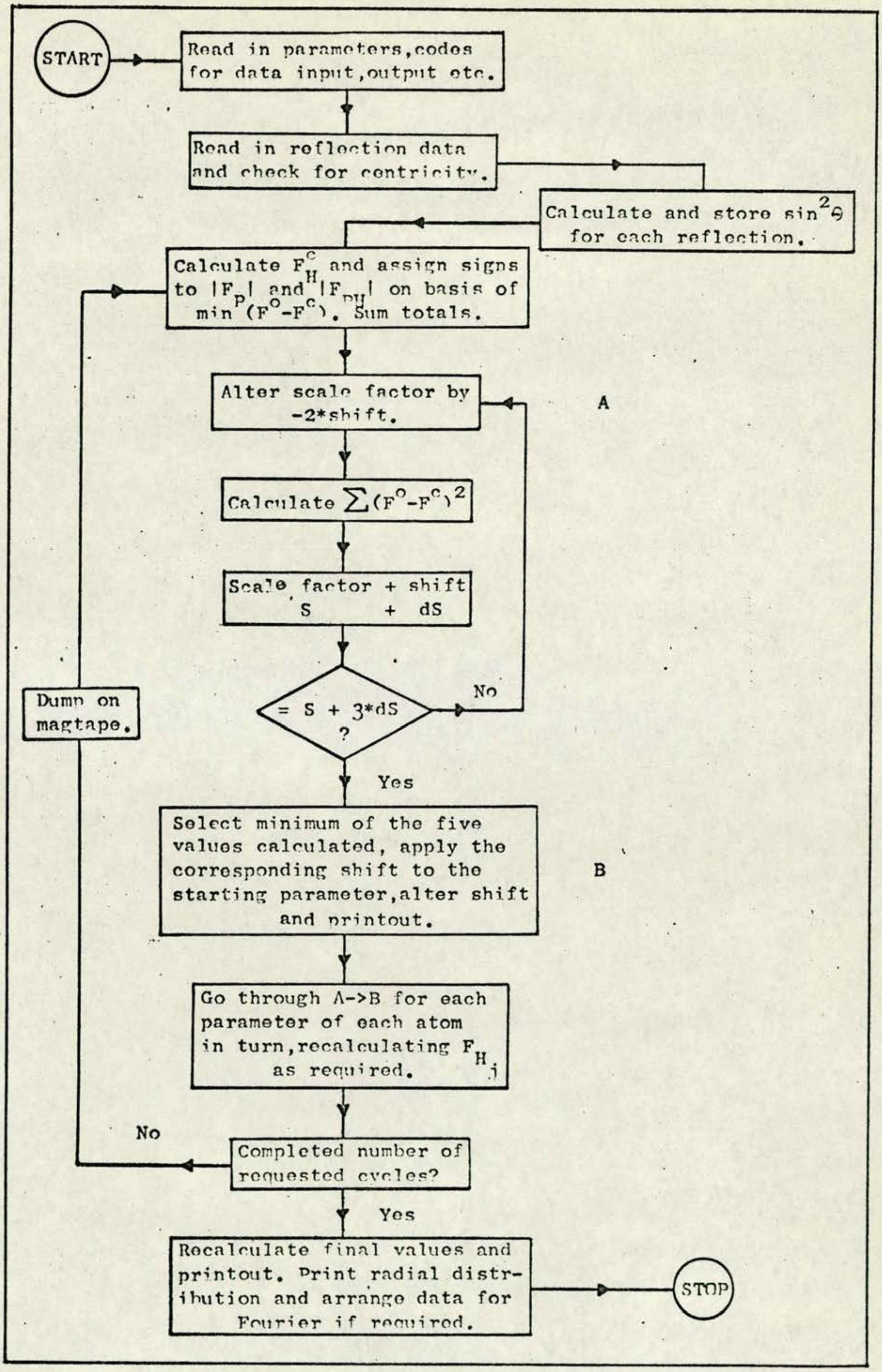


FIG. AI .2 Flow diagram for PANGLOSS.

Initially, then, the program was coded in AA for the KDF 9. Fig. AI. 2 shows the flow diagram for the basic program, the main features of which will be discussed in more detail. The first input was on cards (or paper tape) and included a series of code numbers dealing with such things as which space group symmetry was required for the structure factor calculation, the methods used for input and output of the reflection data, the number of heavy atom sites and the number of cycles of refinement. Also, the unit cell parameters were read in followed by the scale factor and its shift and then the parameters and their shifts for each of the heavy atom sites, in the order occupancy, temperature factor, x,y and z coordinates. These were followed by the coefficients A,. B, C, D for the scattering factor curves for the atoms which were calculated from the expression shown below.

$$f(\sin\theta/\lambda) = A + B.\sin^2\theta/\lambda^2 + C.\sin^4\theta/\lambda^4 + D.\sin^6\theta/\lambda^6$$

Usually the reflection data were read into the store from magtape in the form of lists produced by COLLATOR, each of which has a short alphanumeric string as an identifier. If this sort of data was being input, then the identifier string had also to be fed in.

Data input is in the form containing h,k,l mod(F_{PH}) and mod(F_P). On the KDF 9 routines were available to take this input from the five-track paper tape produced by the Elliott 803 computer at the Royal Institution, from magtape and from cards. On input, each reflection was checked for its centricity relevant to the space group being employed and then packed into one 48 bit word with the signed indices occupying 7 bits each and the structure factors' amplitudes, 13 bits each. The packing and unpacking could be done most efficiently with a small machine-code patch each time this had to be done.

The first calculation was that of the values of $\sin^2\theta$ for each reflection, which were stored in an array parallel to the reflection data. These were the preliminaries which occupied less than a minute of computing time.

Each cycle of refinement was then carried out in the same fashion from this point on. First, the heavy atom contribution from each site was calculated and stored. The total extra contribution from the derivative can then be found by addition and the signs of the $|F_{PH}|$ and $|F_P|$ determined on the basis of which pair gave the smallest difference between observed and calculated. Also, an R factor, defined as

$$R = \frac{\sum_{hkl} (\,|\,|F_H^{obs}| - |F_H^{calc}|\,|)}{\sum_{hkl} |\,F_H^{obs}|}$$

and the total squared differences between observed and calculated were printed out, after the summation was complete.

Next, the scale factor was refined by reducing the value by twice the specified shift and calculating the sum of the squared differences. This value was stored whilst the parameter is increased by its shift and the summing repeated. When 2 times the shift had been applied (5 summations in all), the value of the scale factor giving the smallest value of (Δ^2) was chosen for the next cycle and the shift altered for this next cycle by doubling, halving or quartering the value depending on whether the minimum occurred when ± 2, ± 1 or 0 times the shift had been added to the parameter.

Having completed the scale factor refinement for the current cycle (which is common to all sites), each parameter was then optimised in turn. The only difference was that the heavy atom contribution for the parameter being dealt with had to be recalculated each time. It was found to save a great deal of time by storing the contributions and only calculating that which was affected by the parameter variation. At the end of each completed cycle of refinement the whole store was written to magtape so that it could be picked up from there if, for some reason like the time running out, the operators stopped the job.

If the R factor did not move down within three consecutive cycles, refinement was stopped anyway. At this point, a recalculation was made using the final parameters and all the data reflection, including the calculated heavy atom contributions and the assigned signs, were printed out. Finally, the data could be passed to the Fourier program for calculation of a difference or "double-difference" Fourier projection using the calculated signs.

On the 360/50 the program is almost identical save that the storing of the reflection data in the computer has had to be reorganised on account of the different word length. The data is now held with two words for each reflection, the first holding h,k and l with each index in one byte, the second holding the structure amplitudes in short, 16-bit integers. Because the byte is incapable of containing signed integers between -255 and 255, 128 has to be added to each index before packing.it and then, of course, removed before use. The packing and unpacking of the data is very quick using the compilers built-in addressing routines.

At present, the program only deals with space groups $P3_221$ and $B22_12$ both of which allow the refinement of all three positional parameters. Adaptation to other centrosymmetric space groups would be straightforward since only one routine need be altered. Calculation of the four constants for the scattering factor curve is done using a standard curve fitting program on the PDP-8 computer in this Department.

One problem for which no satisfactory conclusion was reached whilst writing the program was that of whether the alteration of the parameters should be done as each is refined or altogether at the end of the cycle. Tests showed that it made little or no difference in the number of cycles required to refine a given site. Perhaps the only reason for <u>not</u> changing them in the middle of the cycle is if a very large shift occurs which affects all other parameters still to be refined. Since the shifts are usually kept small this has never been found to create a problem. A trivially small amount of extra storage space is required for the end-shift version.

The program was tested initially by working out by hand some structure factors and checking that these caused the appropriate assignations to be made. It was then compared with data run on the Elliott 803 at the Royal Institution and found to give almost identical figures throughout, the differences being due to rounding-errors.

As with all computing problems, a compromise had to be made between the time taken and the space occupied in the machine. How much time can be saved by storing parts of the calculation and yet not running the risk of overflowing the machine's capacity? The first few trials of PANGLOSS were pitifully slow especially with more than one heavy atom. This was because the structure factor routine was very inefficient. Storing all the contributions and only recalculating the one which was changing, helped considerably but still the time was rather too long for each cycle. A chance remark about the time overhead for a routine call alone (i.e. with no calculation) and with no parameters being 200 microseconds on the KDF 9 (an addition taking 1-2 microseconds) brought home one major cause of inefficiency in the structure factor routine: cosine and sine calls were being made over and over again inside the innermost loops. Also, the accuracy of the standard routines was far too great for the purpose here. Consequently, a machine-code patch employing a quickly converging series was used instead. For $x^2 \leq 1$ then

$$\sin \frac{\pi}{2} x = x(1.570795-0.645925x^2+0.079500x^4-0.004370x^6)$$

and

$$\cos \frac{\pi}{2} x = \sin \frac{\pi}{2}(1-x)$$

The error is less than 1×10^{-5}. This gave an improvement of a factor of 3 in the time taken for each cycle. This was only applicable to the KDF 9 and on the 360/50 no such patch has been employed. It is probable, however, that a look-up table with some crude interpolation procedure would speed the program up markedly.

Number of Atoms	Number of Parameters	Number of Reflections	Time (secs)
3	13	139	132
3	6	139	75
2	8	144	89
2	8	185	120
1	4	185	49
1	1	185	23

TABLE AI.1 Typical times per cycle of the refinement programme, PANGLOSS.

Currently the program is held on disc in a compiled form and can be easily retrieved for use. Typical times for operation are in Table AI. 1.

APPENDIX II

The Rate of Reaction of TCA and β-Lactoglobulin

Experimental.

100mls. of a 1.0mM stock protein solution were made up in 0.3M mixed phosphate buffer at pH about 6.0. About 10mls. of this solution were stored in a suitably labelled bottle at $3^{o}C$ and the remainder of the solution carefully titrated to pH 6.5 with 4M NaOH. A small portion was again removed and stored. This process was repeated another six times until pH 8.5 was reached. In addition, 4M solutions of NaCl, A.S. and mixed phosphate were made up and taken to pH 8.15, measured on samples diluted tenfold.

The reactions were carried out in a 10mm silica cuvette in the thermostatted cell holder of the Unicam SP500 spectrophotometer. The temperature was maintained at $24.8 \pm 0.1^{o}C$ by means of a water bath and pump. The blank was 0.1M phosphate and the order of addition of the solution was:

2.00mls. distilled water

1.00mls. protein solution at each pH

These were mixed and the optical density at 245nm and 280nm measured. Measurement of the absorbance at 280nm was used to check the concentration of the protein given that $E^{1\%}_{280} = 9.8$ (Pantaloni (1965)).

0.05mls. stock 0.025M stock TCA solution

This final addition was made using a stirrer with a flattened, upturned end (a "plumper") onto which the TCA was pipetted. This was then plunged into the cuvette at the same time as the stop-clock was started. Readings were made at 245nm at time intervals which were adjusted to give a change in absorbance of at least 0.01. It was found that the reaction gave a final optical density reading of about 0.55 which is in a range of the instrument where the scale divisions correspond to 0.02 absorbance units. Because of this, the readings taken above 0.5 O.D. (or about 85% reaction) were disregarded in the calculation of the rates. The final reading was unfortunately in this range so that the average of about a dozen readings was taken as the true end point of the reaction. This measurement was repeated some 24 hours later in order to check that the reaction had, in fact, reached completion. In the case of runs near pH 7.0 the final value was checked a third time after a further 24 hours.

Solvation measurements were made in an analogous manner to those for the rate except that the protein aliquot was replaced by one of 0.3M phosphate buffer.

The reactions in the presence of salt were exactly as for the salt-free only 0.75mls of salt solution replaced the same amount of distilled water. References were of the same concentration as the salt used and solvation measurements were made as before.

Calculation.

The second order rate equation can be written as

$$t.k_{obs} + c = \frac{2.303}{(a-b)} \;.\log \left(\frac{b}{a.} \frac{(a-x)}{(b-x)} \right) \text{ (e.g. see Moore (1963))}$$

where at time, t, x moles/l. of product will have been formed from the initial concentrations in moles/l., a and b, of the two reactants. If b is the minor component, then y, the fraction reacted at time t will be given by

$$y.b = x.$$

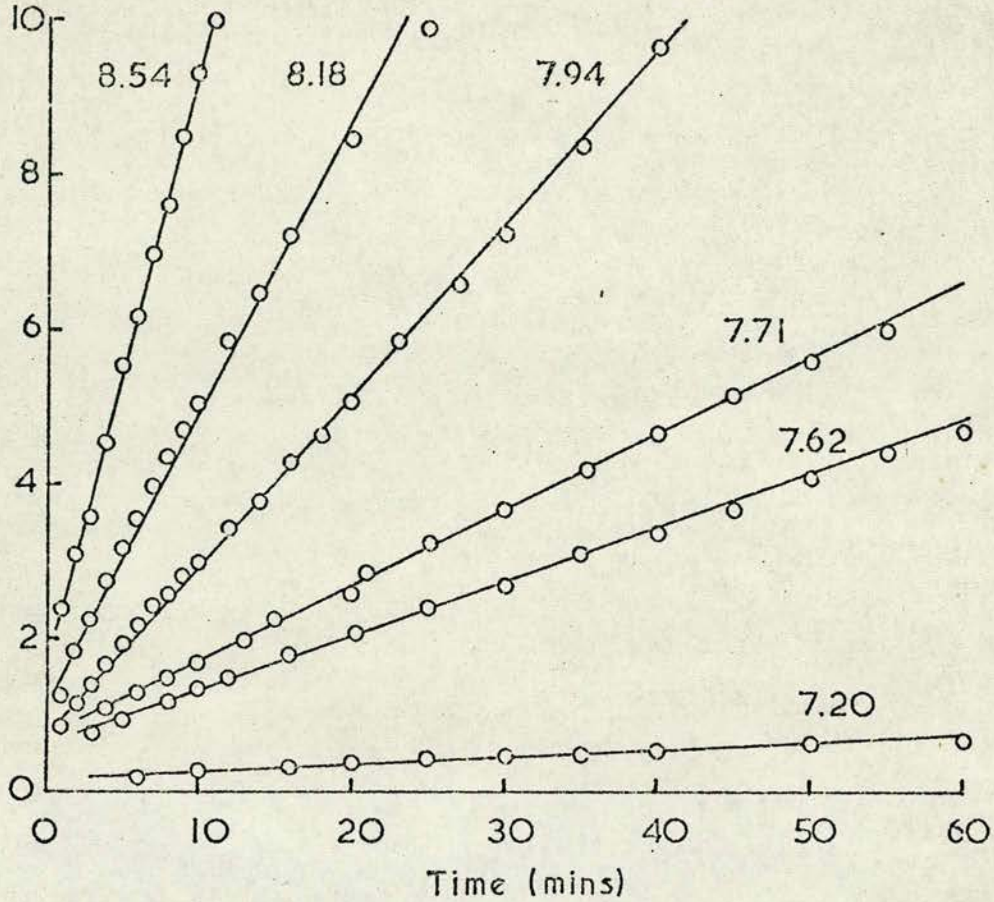


FIG. AII.1 Plot of 2.303/(a-b).log(1-by/a)/(1-y) - c versus time for the rate of reaction of TCA with lactoglobulin at various pH values.

Substituting then gives

$$t.k_{obs} + c = \frac{2.303}{(a-b)} .\log \left(\frac{(1-b/a.y)}{(1-y)} \right).$$

A plot of the righthand side of this equation against t will then be a straight line of slope k_{obs} and intercept at t=0, c. The fraction y, in this case was given by

$$y = \frac{E_t - E_i}{E_f - E_i}$$

where E was the optical density at 245nm and the subscript referred to the initial reading at t=0, i, the reading at time, t, and the final reading on completion of the reaction, f.

Basolo and Pearson (1967) have expressed the rate of substitution of square planar complexes as

$$k_{obs} = (k_1 + k_2.[\text{substituent}]).[\text{complex}]$$

where k_1 refers to the solvation of the complex. However, Cattalini and his coworkers have shown that k_1 is small enough to be ignored in the reactions of Au(III) complexes (Cattalini et al., (1968)). The constant term, c, is an extra term which was introduced in this case and must refer to a very fast, solvation-type reaction which is not concerned with the substitution at the sulphydryl group since it is practically unaffected by either salt or pH.

The data for the reactions followed are shown overleaf and Fig. AII. 1 is a plot of one pair of measurements showing the righthand side of equation (1) plotted against t.

SERIES OF RUNS OF PH = 7.46 WITH 0.1M PHOSPHATE

SECOND ORDER RATE CALCULATION FOR SALT FREE RUN 1

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002363 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.223
FINAL OPTICAL DENSITY READING = 0.546

TIME(SECS)	O.D.	%REACTION	LOG TERM
30	0.247	7.43	434.6
60	0.253	9.29	549.1
120	0.259	11.15	666.0
180	0.263	12.38	745.3
240	0.267	13.62	825.9
300	0.271	14.86	907.7
360	0.277	16.72	1032.7
480	0.282	18.27	1139.2
600	0.290	20.74	1314.0
720	0.298	23.22	1494.6
960	0.308	26.32	1729.2
1215	0.321	30.34	2049.9
1500	0.333	34.06	2363.6
1800	0.344	37.46	2667.7
2100	0.358	41.80	3080.6
2160	0.358	41.80	3080.6
2400	0.366	44.27	3331.1
2700	0.376	47.37	3660.9
3000	0.386	50.46	4011.3
3300	0.396	53.56	4385.0
3600	0.403	55.73	4662.2
4260	0.420	60.99	5397.9
4800	0.431	64.40	5930.3
5400	0.442	67.80	6517.4
6000	0.453	71.21	7171.6
6600	0.461	73.68	7698.8
12120	0.497	84.83	10942.0

VALUE OF K2 = 9.49601a -1 LITRES/MOLE/SEC.

INTERCEPT = 3.17693a +3

SERIES OF RUNS OF PH = 8.51 WITH 0.1M PHOSPHATE

SECOND ORDER RATE CALCULATION FOR SALT FREE RUN 2

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002512 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.267
FINAL OPTICAL DENSITY READING = 0.540

TIME(SECS)	O.D.	%REACTION	LOG TERM
45	0.336	25.27	1650.0
75	0.357	32.97	2271.3
105	0.373	38.83	2796.8
135	0.388	44.32	3339.3
165	0.399	48.35	3773.6
195	0.411	52.75	4289.2
225	0.420	56.04	4709.5
255	0.429	59.34	5163.4
285	0.437	62.27	5599.9
330	0.448	66.30	6260.3
390	0.459	70.33	7006.9
450	0.468	73.63	7699.0
510	0.477	76.92	8485.4
570	0.483	79.12	9075.8
630	0.489	81.32	9732.9
690	0.495	83.52	10473.3

VALUE OF K2 = 1.37348a +1 LITRES/MOLE/SEC.

INTERCEPT = 1.61344a +3

SERIES OF RUNS OF PH = 7.80 WITH 0.1M PHOSPHATE

SECOND ORDER RATE CALCULATION FOR SALT FREE RUN 3

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002506 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.250
FINAL OPTICAL DENSITY READING = 0.540

TIME(SECS)	O.D.	%REACTION	LOG TERM
30	0.285	12.07	725.2
60	0.291	14.14	860.0
90	0.297	16.21	998.3
120	0.302	17.93	1116.3
180	0.311	21.03	1335.5
240	0.320	24.14	1563.9
300	0.328	26.90	1775.2
420	0.344	32.41	2224.2
540	0.359	37.59	2681.2
660	0.370	41.38	3047.1
780	0.382	45.52	3464.5
900	0.394	49.66	3921.5
1020	0.403	52.76	4290.4
1140	0.412	55.86	4685.2
1470	0.434	63.45	5784.8
1740	0.450	68.97	6742.7
2250	0.472	76.55	8390.7

VALUE OF K2 = 3.41658a +0 LITRES/MOLE/SEC.

INTERCEPT = 2.62612a +3

SERIES OF RUNS OF PH = 7.60 WITH 0.1M PHOSPHATE

SECOND ORDER RATE CALCULATION FOR SALT FREE RUN 2

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002386 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.229
FINAL OPTICAL DENSITY READING = 0.540

TIME(SECS)	O.D.	%REACTION	LOG TERM
30	0.260	9.97	591.6
60	0.263	10.93	652.4
120	0.270	13.18	797.2
180	0.277	15.43	946.0
240	0.283	17.36	1076.9
240	0.286	18.33	1143.6
360	0.290	19.61	1233.7
480.	0.300	22.83	1465.9
600	0.309	25.72	1683.7
780	0.321	29.58	1988.2
900	0.329	32.15	2201.1
1020	0.337	34.73	2422.5
1140	0.343	36.66	2594.6
1260	0.350	38.91	2802.4
1500	0.363	43.09	3210.2
1800	0.375	46.95	3615.2
2130	0.389	51.45	4127.9
2400	0.401	55.31	4608.0
2700	0.413	59.16	5132.6
3000	0.427	62.06	5560.6
3300	0.430	64.63	5970.0
3900	0.446	69.77	6888.9
4500	0.459	73.95	7761.6
5160	0.468	76.85	8453.5
5700	0.474	78.78	8965.5
6300	0.480	80.71	9526.9
6900	0.486	82.64	10148.3

VALUE OF K2 = 1.17568a +0 LITRES/MOLE/SEC.

INTERCEPT = 3.24085a +3

SERIES OF RUNS OF PH = 8.51 WITH 0.1M PHOSPHATE

SECOND ORDER RATE CALCULATION FOR SALT FREE RUN 1

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002449 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.260
FINAL OPTICAL DENSITY READING = 0.542

TIME(SECS)	O.D.	%REACTION	LOG TERM
30	0.328	24.11	1561.7
60	0.351	32.27	2211.4
90	0.368	38.30	2746.3
120	0.384	43.97	3301.8
150	0.396	48.23	3758.1
180	0.411	53.55	4386.1
210	0.422	57.45	4895.7
240	0.430	60.28	5297.3
270	0.438	63.12	5729.6
300	0.445	65.60	6136.6
330	0.453	68.44	6640.3
360	0.458	70.21	6979.1
420	0.468	73.76	7723.1
480	0.477	76.95	8485.8
540	0.485	79.79	9259.8
600	0.490	81.56	9801.6
660	0.497	84.04	10656.0

VALUE OF K2 = 1.41569a +1 LITRES/MOLE/SEC.

INTERCEPT = 1.65997a +3

SERIES OF RUNS OF PH = 7.00 WITH 0.1M PHOSPHATE

SECOND ORDER RATE CALCULATION FOR SALT FREE RUN 2

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002633 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.259
FINAL OPTICAL DENSITY READING = 0.575

TIME(SECS)	O.D.	%REACTION	LOG TERM
60	0.264	1.58	89.6
270	0.268	2.85	162.5
780	0.272	4.11	236.3
2520	0.286	8.54	503.1
3840	0.293	10.76	641.7
4320	0.296	11.71	702.2
4800	0.300	12.97	784.0
5280	0.302	13.61	825.3
6480	0.307	15.19	930.1
11820	0.332	23.10	1487.1
13260	0.339	25.32	1653.8
17100	0.355	30.38	2055.4
22860	0.377	37.34	2660.4
27120	0.390	41.46	3051.8
70200	0.482	70.57	7063.7
73200	0.488	72.47	7450.3
80400	0.500	76.27	8331.9

VALUE OF K2 = 8.16512a -2 LITRES/MOLE/SEC.

INTERCEPT = 3.75922a +3

SERIES OF RUNS OF PH = 7.80 WITH 0.1M PHOSPHATE

SECOND ORDER RATE CALCULATION FOR SALT FREE RUN 2

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002472 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.250
FINAL OPTICAL DENSITY READING = 0.530

TIME(SECS)	O.D.	%REACTION	LOG TERM
30	0.284	12.14	729.9
60	0.290	14.29	869.7
90	0.296	16.43	1013.3
120	0.301	18.21	1135.9
180	0.310	21.43	1363.8
240	0.320	25.00	1628.8
315	0.330	28.57	1907.4
360	0.336	30.71	2081.6
420	0.343	33.21	2292.1
480	0.351	36.07	2542.9
540	0.358	38.57	2772.2
600	0.363	40.36	2941.9
720	0.375	44.64	3371.8
840	0.385	48.21	3757.2
960	0.397	52.50	4257.7
1080	0.405	55.36	4617.8
1200	0.414	58.57	5052.5
1620	0.440	67.86	6534.7
1800	0.450	71.43	7225.3
2100	0.464	76.43	8356.4
2400	0.477	81.07	9650.1
2700	0.483	83.21	10360.1

VALUE OF K2 = 3.50165a +0 LITRES/MOLE/SEC.

INTERCEPT = 2.81114a +3

SERIES OF RUNS OF PH = 7.00 WITH 0.1M PHOSPHATE

SECOND ORDER RATE CALCULATION FOR SALT FREE RUN 1

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002673 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.264
FINAL OPTICAL DENSITY READING = 0.580

TIME(SECS)	O.D.	%REACTION	LOG TERM
30	0.268	1.27	71.6
60	0.269	1.58	89.6
120	0.269	1.58	89.6
240	0.273	2.85	162.5
360	0.275	3.48	199.2
600	0.277	4.11	236.3
960	0.281	5.38	311.2
1200	0.284	6.33	368.0
1500	0.286	6.96	406.3
1800	0.289	7.91	464.2
2100	0.290	8.23	483.6
2400	0.292	8.86	522.7
3000	0.296	10.13	601.8
3600	0.300	11.39	682.0
4440	0.306	13.29	804.7

VALUE OF K2 = 8.21524a -1 LITRES/MOLE/SEC.

INTERCEPT = 2.90068a +3

SERIES OF RUNS OF PH = 8.15 WITH 0.1M PHOSPHATE

SECOND ORDER RATE CALCULATION FOR SALT FREE RUN 2

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002541 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.267
FINAL OPTICAL DENSITY READING = 0.571

TIME(SECS)	O.D.	%REACTION	LOG TERM
50	0.330	20.72	1313.2
75	0.342	24.67	1604.2
105	0.352	27.96	1859.2
135	0.363	31.58	2154.1
165	0.372	34.54	2407.8
195	0.381	37.50	2673.7
225	0.390	40.46	2953.0
255	0.398	43.09	3213.7
300	0.407	46.05	3522.3
360	0.419	50.00	3962.3
420	0.431	53.95	4439.6
480	0.440	56.91	4826.1
540	0.451	60.53	5337.4
600	0.459	63.16	5740.4
660	0.465	65.13	6062.4
780	0.479	69.74	6892.8
900	0.490	73.36	7641.3
1020	0.502	77.30	8586.1
1140	0.507	78.95	9030.2
1260	0.516	81.91	9926.2
1440	0.523	84.21	10732.3

VALUE OF K2 = 6.24821a +0 LITRES/MOLE/SEC.

INTERCEPT = 2.65792a +3

SERIES OF RUNS OF PH = 8.51 WITH 0.1M PHOSPHATE

SECOND ORDER RATE CALCULATION FOR SALT FREE RUN 3

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002644 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.070
FINAL OPTICAL DENSITY READING = 1.410

TIME(SECS)	O.D.	%REACTION	LOG TERM
60	0.530	34.33	2390.4
120	0.630	41.79	3085.1
180	0.690	46.27	3547.9
240	0.800	54.48	4510.5
360	0.950	65.67	6160.0
420	1.010	70.15	6981.1
480	1.050	73.13	7601.6
540	1.100	76.87	8484.3
600	1.140	79.85	9301.6
660	1.170	82.09	9999.7
720	1.190	83.58	10516.0

VALUE OF K2 = 1.32591a +1 LITRES/MOLE/SEC.

INTERCEPT = 1.64884a +3

SERIES OF RUNS OF PH = 7.80 WITH 0.1M PHOSPHATE

SECOND ORDER RATE CALCULATION FOR SALT FREE RUN 1

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002506 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.260
FINAL OPTICAL DENSITY READING = 0.550

TIME(SECS)	O.D.	%REACTION	LOG TERM
30	0.293	11.38	681.0
60	0.307	14.48	882.8
90	0.308	16.55	1021.7
120	0.313	18.28	1140.3
150	0.318	20.00	1261.4
180	0.324	22.07	1410.5
210	0.329	23.79	1538.0
240	0.333	25.17	1642.2
270	0.338	26.90	1775.2
300	0.343	28.62	1911.6
360	0.352	31.72	2166.0
420	0.360	34.48	2402.4
480	0.367	36.90	2618.0
540	0.373	38.97	2809.7
600	0.380	41.38	3042.1
720	0.392	45.52	3464.5
840	0.404	49.66	3921.5
960	0.414	53.10	4333.0
1080	0.423	56.21	4730.8
1200	0.431	58.97	5109.7
1380	0.445	63.79	5840.2
1560	0.456	67.59	6487.8
1740	0.466	71.03	7147.6
2100	0.482	76.55	8390.7
2415	0.495	81.03	9642.9
2700	0.503	83.79	10572.7

VALUE OF K2 = 3.94782a +0 LITRES/MOLE/SEC.

INTERCEPT = 2.95869a +3

SERIES OF RUNS OF PH = 8.15 WITH 0.1M PHOSPHATE

SECOND ORDER RATE CALCULATION FOR SALT FREE RUN 1

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002541 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.272
FINAL OPTICAL DENSITY READING = 0.578

TIME(SECS)	O.D.	%REACTION	LOG TERM
30	0.318	15.03	919.5
60	0.333	19.93	1256.9
90	0.344	23.53	1518.5
120	0.356	27.45	1818.9
150	0.364	30.07	2028.8
180	0.371	32.35	2219.3
210	0.380	35.29	2474.4
240	0.388	37.91	2711.4
270	0.392	39.22	2833.9
300	0.402	42.48	3152.3
360	0.413	46.08	3525.1
420	0.425	50.00	3962.3
480	0.434	52.94	4314.0
540	0.443	55.88	4689.2
600	0.451	58.50	5044.9
720	0.467	63.73	5831.1
840	0.478	67.32	6442.2
960	0.490	71.24	7192.5
1080	0.500	74.51	7902.1
1200	0.507	76.80	8456.3
1350	0.512	78.43	8887.2
1500	0.522	81.70	9858.3
1560	0.525	82.68	10184.2

VALUE OF K2 = 6.27351a +0 LITRES/MOLE/SEC.

INTERCEPT = 2.95079a +3

SERIES OF RUNS OF PH=8.15 WITH 1M SODIUM CHLORIDE

SECOND ORDER RATE CALCULATION FOR 1 MOLAR CHLORIDE RUN 2

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002909 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.304
FINAL OPTICAL DENSITY READING = 0.580

TIME(SECS)	O.D.	%REACTION	LOG TERM
240	0.340	13.04	788.8
360	0.342	13.77	836.3
2160	0.356	18.84	1180.8
3600	0.365	22.10	1414.5
4680	0.372	24.64	1603.6
6000	0.380	27.54	1828.1
11340	0.408	37.68	2695.4
12720	0.413	39.49	2865.8
14580	0.422	42.75	3186.3
17100	0.432	46.38	3565.6
18900	0.440	49.28	3888.7
21060	0.448	52.17	4231.5
22980	0.453	53.99	4456.9
26040	0.462	57.25	4886.8
28500	0.469	59.78	5245.2
34860	0.483	64.86	6037.6
37260	0.488	66.67	6349.4
37260	0.488	66.67	6349.4
39900	0.493	68.48	6679.1
34860	0.483	64.86	6037.6
37260	0.488	66.67	6349.4
39900	0.493	68.48	6679.1

VALUE OF K2 = 7.50504a -2 LITRES/MOLE/SEC.

INTERCEPT = 3.37134a +3

SERIES OF RUNS OF PH=8.15 WITH 1M SODIUM CHLORIDE

SECOND ORDER RATE CALCULATION FOR 1 MOLAR CHLORIDE RUN 3

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002552 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.269
FINAL OPTICAL DENSITY READING = 0.505

TIME(SECS)	O.D.	%REACTION	LOG TERM
30	0.328	25.00	1629.3
60	0.337	28.81	1927.4
90	0.343	31.36	2135.6
120	0.349	33.90	2352.0
180	0.362	39.41	2852.2
240	0.372	43.64	3270.2
300	0.382	47.88	3722.1
360	0.391	51.69	4162.6
420	0.398	54.66	4530.8
480	0.405	57.63	4924.6
720	0.426	66.53	6302.1
900	0.439	72.03	7357.6
1500	0.467	83.90	10618.0

VALUE OF K2 = 7.42262a -2 LITRES/MOLE/SEC.

INTERCEPT = 3.40303a +3

SERIES OF RUNS OF PH=8.15 WITH 1M AMMONIUM SULPHATE

SECOND ORDER RATE CALCULATION FOR 1 MOLAR SULPHATE RUN 2

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002788 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.283
FINAL OPTICAL DENSITY READING = 0.542

TIME(SECS)	O.D.	%REACTION	LOG TERM
30	0.315	12.36	744.0
60	0.305	8.49	500.1
90	0.307	9.27	548.0
120	0.309	10.04	596.3
180	0.311	10.81	645.1
300	0.318	13.51	819.4
1320	0.361	30.12	2034.9
1380	0.364	31.27	2130.8
2100	0.388	40.54	2964.8
4560	0.446	62.93	5718.1
5580	0.460	68.34	6644.9
5640	0.464	69.88	6939.7
8820	0.493	81.08	9693.2
10860	0.503	84.94	11052.1

VALUE OF K2 = 1.00781a -1 LITRES/MOLE/SEC.

INTERCEPT = 3.68647a +3

SERIES OF RUNS OF PH=8.15 WITH 1M SODIUM CHLORIDE

SECOND ORDER RATE CALCULATION FOR 1 MOLAR CHLORIDE RUN 4

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002713 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.277
FINAL OPTICAL DENSITY READING = 0.508

TIME(SECS)	O.D.	%REACTION	LOG TERM
30	0.333	24.24	1572.7
60	0.340	27.27	1806.0
90	0.346	29.87	2014.2
120	0.350	31.60	2157.6
180	0.359	35.50	2494.5
240	0.370	40.26	2936.4
300	0.378	43.72	3281.4
360	0.385	46.75	3601.9
420	0.392	49.78	3941.9
480	0.399	52.81	4303.9
540	0.405	55.41	4633.7
600	0.410	57.58	4924.1
720	0.420	61.90	5553.6
900	0.432	67.10	6413.8
1200	0.448	74.03	7806.8
1740	0.468	82.68	10209.5

VALUE OF K2 = 7.34631a -2 LITRES/MOLE/SEC.

INTERCEPT = 3.43577a +3

SERIES OF RUNS OF PH=8.15 WITH 1M SODIUM CHLORIDE

SECOND ORDER RATE CALCULATION FOR 1 MOLAR CHLORIDE RUN 1

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002811 MOLFS/L.
INITIAL OPTICAL DENSITY READING = 0.281
FINAL OPTICAL DENSITY READING = 0.555

TIME(SECS)	O.D.	%REACTION	LOG TERM
240	0.314	12.04	723.9
480	0.317	13.14	794.9
1020	0.320	14.23	866.8
2220	0.327	16.79	1038.5
3900	0.333	18.98	1190.2
6600	0.346	23.72	1534.2
8760	0.357	27.74	1843.3
11820	0.370	32.48	2232.8
13440	0.377	35.04	2454.6

VALUE OF K2 = 7.68487a -2 LITRES/MOLE/SEC.

INTERCEPT = 3.42058a +3

SERIES OF RUNS OF PH=8.15 WITH 1M PHOSPHATE BUFFER

SECOND ORDER RATE CALCULATION FOR 1 MOLAR PHOSPHATE RUN 7

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002805 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.279
FINAL OPTICAL DENSITY READING = 0.533

TIME(SECS)	O.D.	%REACTION	LOG TERM
30	0.305	10.24	608.8
60	0.284	1.97	111.8
180	0.288	3.54	202.9
300	0.291	4.72	272.3
1560	0.302	9.06	534.8
3000	0.311	12.60	759.7
4860	0.323	17.32	1075.1
5520	0.327	18.90	1184.5
9900	0.352	28.74	1923.3
13620	0.371	36.22	2560.5
14940	0.377	38.58	2778.0
18000	0.387	42.52	3160.8
16620	0.382	40.55	2966.1
21840	0.400	47.64	3701.3
23640	0.406	50.00	3969.6
36240	0.437	62.20	5604.6
38400	0.441	63.78	5854.4
42000	0.448	66.54	6319.7

VALUE OF K2 = 7.79153a -2 LITRES/MOLE/SEC.

INTERCEPT = 3.47635a +3

SERIES OF RUNS OF PH = 7.60 WITH 0.1M PHOSPHATE

SECOND ORDER RATE CALCULATION FOR SALT FREE RUN 1

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002420 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.242
FINAL OPTICAL DENSITY READING = 0.560

TIME(SECS)	O.D.	%REACTION	LOG TERM
30	0.271	9.12	538.6
60	0.276	10.69	637.2
120	0.282	12.58	758.0
180	0.287	14.15	860.8
240	0.293	16.04	986.7
300	0.297	17.30	1072.3
360	0.303	19.18	1203.4
420	0.307	20.44	1292.5
480	0.311	21.70	1383.2
600	0.320	24.53	1592.8
720	0.329	27.36	1810.9
840	0.336	29.56	1986.7
960	0.343	31.76	2168.3
1080	0.350	33.96	2356.1
1200	0.357	36.16	2550.6
1500	0.372	40.88	2991.8
1800	0.387	45.60	3471.1
2100	0.400	49.69	3922.6
2400	0.411	53.14	4335.3
2700	0.422	56.60	4780.5
3000	0.432	59.75	5218.1
3300	0.441	62.58	5642.9
3600	0.450	65.41	6102.0
3900	0.457	67.61	6486.5
4200	0.465	70.13	6959.9
4800	0.474	72.96	7543.6
5400	0.483	75.79	8193.0
6000	0.490	77.99	8753.9
6600	0.497	80.19	9374.7
7200	0.504	82.39	10069.7

VALUE OF K2 = 1.29785a +0 LITRES/MOLE/SEC.

INTERCEPT = 3.40194a +3

SERIES OF RUNS OF PH = 7.46 WITH 0.1M PHOSPHATE

SECOND ORDER RATE CALCULATION FOR SALT FREE RUN 2

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002368 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.218
FINAL OPTICAL DENSITY READING = 0.540

TIME(SECS)	O.D.	%REACTION	LOG TERM
30	0.242	7.45	436.0
120	0.252	10.56	628.8
60	0.245	8.39	493.1
180	0.256	11.80	707.9
240	0.262	13.66	828.7
360	0.268	15.53	952.3
420	0.273	17.08	1057.5
480	0.277	18.32	1143.2
600	0.284	20.50	1296.4
930	0.300	25.47	1663.8
1260	0.318	31.06	2109.0
1560	0.330	34.78	2427.7
1800	0.340	37.89	2707.2
2040	0.349	40.68	2971.8
2280	0.359	43.79	3281.4
2580	0.370	47.20	3643.1
2820	0.376	49.07	3850.7
3120	0.384	51.55	4140.1
3420	0.394	54.66	4524.0
3900	0.408	59.01	5109.4
4500	0.419	62.42	5616.0
5100	0.430	65.84	6171.9
6300	0.449	71.74	7281.2
6900	0.457	74.22	7820.9
7500	0.465	76.71	8416.1
8100	0.470	78.26	8821.8
8700	0.475	79.81	9258.0
9420	0.481	81.68	9828.7
10920	0.488	83.85	10573.7

VALUE OF K2 = 8.42641a -1 LITRES/MOLE/SEC.

INTERCEPT = 3.08297a +3

SERIES OF RUNS OF PH=8.15 WITH 1M PHOSPHATE BUFFER

SECOND ORDER RATE CALCULATION FOR 1 MOLAR PHOSPHATE RUN 1

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002823 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.289
FINAL OPTICAL DENSITY READING = 0.533

TIME(SECS)	O.D.	%REACTION	LOG TERM
30	0.299	4.10	235.4
60	0.294	2.05	116.4
120	0.297	3.28	187.5
240	0.301	4.92	283.8
300	0.302	5.33	308.1
600	0.306	6.97	406.6
960	0.308	7.79	456.6
1500	0.311	9.02	532.4
2460	0.317	11.48	687.4
3900	0.325	14.75	901.4
5340	0.331	17.21	1067.6
10260	0.356	27.46	1821.4
12300	0.364	30.74	2086.5
14640	0.374	34.84	2437.0
16680	0.382	38.11	2734.5
18420	0.388	40.57	2968.6
20340	0.393	42.62	3171.5
22260	0.398	44.67	3382.1
61440	0.463	71.31	7228.9

VALUE OF K2 = 8.63004a -2 LITRES/MOLE/SEC.

INTERCEPT = 3.56198a +3

SERIES OF RUNS OF PH=8.15 WITH 1M AMMONIUM SULPHATE

SECOND ORDER RATE CALCULATION FOR 1 MOLAR SULPHATE RUN 1.

INITIAL CONCENTRATION OF LIGAND = 0.00040996 MOLES/L.
INITIAL CONCENTRATION OF PROTEIN = 0.00002771 MOLES/L.
INITIAL OPTICAL DENSITY READING = 0.283
FINAL OPTICAL DENSITY READING = 0.550

TIME(SECS)	O.D.	%REACTION	LOG TERM
30	0.322	14.61	891.5
60	0.304	7.87	461.4
90	0.308	9.36	554.0
120	0.309	9.74	577.4
180	0.311	10.49	624.6
240	0.315	11.99	720.1
300	0.318	13.11	792.9
600	0.335	19.48	1225.1
960	0.350	25.09	1637.6
1380	0.368	31.84	2177.7
1620	0.379	35.96	2536.1
3420	0.431	55.43	4638.4
5100	0.464	67.79	6542.4
6300	0.482	74.53	7928.3
7080	0.485	75.66	8195.3
7920	0.495	79.40	9185.6
8940	0.500	81.27	9751.7
10140	0.507	83.90	10649.0

VALUE OF K2 = 9.26898a -2 LITRES/MOLE/SEC.

INTERCEPT = 3.72180a +3

APPENDIX III

Reflection Data to 6Å Resolution for the

TCA and TCP Derivatives

F_{PH1} : TCA

F_{PH2} : TCP

	H	K	L	|FPH1|	|FPH2|	|FP|	PHASE	
*	0	0	2	1168		743	180	0
*	0	0	4	733		77	0	0
*	0	0	6	877	1525	1316	0	0
*	0	0	8	1504	1272	1768	180	0
*	0	0	10	960	259	1319	0	0
*	0	0	12	168	457	668	180	0
*	0	1	2	380		310	90	0
*	0	1	4	448	215	411	90	0
*	0	1	6	78	570	264	270	0
*	0	1	8	312	2083	1169	270	0
*	0	1	10	792	1671	1286	270	0
*	0	1	12	182	696	472	270	0
*	0	2	0	626		1535	0	0
*	0	2	2	725	67	216	180	0
*	0	2	4	1967	1310	1334	0	0
*	0	2	6	501	596	864	0	0
*	0	2	8	1562	1459	1931	180	0
*	0	2	10	423	89	610	0	0
*	0	2	12	777	623	705	0	0
*	0	3	2	299	549	915	90	0
*	0	3	4	99	100	149	270	0
*	0	3	6	604	460	293	90	0
*	0	3	8	114	76	152	270	0
*	0	3	10	678	379	415	270	0
*	0	3	12	601	115	119	90	0
*	0	4	0	1019		1565	0	0
*	0	4	2	464	78	43	180	0
*	0	4	4	172	667	844	180	0
*	0	4	6	231	499	796	0	0
*	0	4	8	518	718	514	0	0
*	0	4	10	218	419	567	180	0
*	0	4	12	174	110	235	180	0
*	0	5	2	938	762	1347	270	0
*	0	5	4	891	1085	1257	270	0
*	0	5	6	337	678	218	270	0
*	0	5	8	1019	417	644	90	0
*	0	5	10	301	197	72	270	0
*	0	5	12	211	109	247	270	0
*	0	6	0	140	114	118	0	0
*	0	6	2	669	469	242	180	0
*	0	6	4	1067	660	783	180	0
*	0	6	6	199	111	124	0	0
*	0	6	8	889	559	877	0	0
*	0	6	10	146	150	244	180	0
*	0	7	2	1223	1232	1078	90	0
*	0	7	4	291	644	687	270	0
*	0	7	6	525	68	315	270	0
*	0	7	8	133	412	70	270	0

	H	K	L	|FPH1|	|FPH2|	|FP|	PHASE	
*	0	7	10	139	168	139	90	0
*	0	8	0	1064	595	310	0	0
*	0	8	2	776	400	693	180	0
*	0	8	4	226	332	200	180	0
*	0	8	6	216	78	406	0	0
*	0	8	8	455	480	730	180	0
*	0	9	2	584	516	659	270	0
*	0	9	4	191	63	124	90	0
*	0	9	6	357	230	80	90	0
*	0	9	8	332	772	655	90	0
*	0	10	0	977	839	294	180	0
*	0	10	2	368	83	99	180	0
*	0	10	4	744	572	618	0	0
*	0	10	6	146	406	90	180	0
*	0	11	2	214	226	96	90	0
*	2	0	0	124	662	72	0	0
*	2	0	2	687	913	767	180	0
*	2	0	4	780	549	1328	180	0
*	2	0	6	957	204	763	180	0
*	2	0	8	1120	1219	1461	0	0
*	2	0	10	393	1243	981	180	0
*	2	0	12	185	336	77	180	0
*	2	1	0	212	108	216	180	0
*	2	1	2	609	872	643	96	0
*	2	1	4	608	243	427	339	0
*	2	1	6	564	527	872	202	3
*	2	1	8	262	338	415	121	1
*	2	1	10	63	392	333	314	0
*	2	1	12	365	141	430	121	37
*	2	2	0	843	574	807	180	0
*	2	2	2	508	327	384	249	31
*	2	2	4	1162	829	1055	85	28
*	2	2	6	918	653	725	306	15
*	2	2	8	802	507	711	184	38
*	2	2	10	154	145	112	243	26
*	2	2	12	287	394	415	59	8
*	2	3	0	164	432	183	180	0
*	2	3	2	254	773	460	294	19
*	2	3	4	1237	855	672	243	0
*	2	3	6	506	692	248	102	0
*	2	3	8	824	1042	1087	75	14
*	2	3	10	301	129	400	254	0
*	2	3	12	561	378	339	172	29
*	2	4	0	1358	1357	1496	180	0
*	2	4	2	568	285	663	81	1
*	2	4	4	1046	692	1241	255	37
*	2	4	6	373	272	63	212	0
*	2	4	8	587	727	641	29	36

	H	K	L	|FPH1|	|FPH2|	|FP|	PHASE	
*	2	4	10	718	988	966	155	0
*	2	4	12	237	68	231	73	18
*	2	5	0	803	256	442	0	0
*	2	5	2	874	660	604	237	0
*	2	5	4	556	585	765	106	0
*	2	5	6	268	437	317	95	12
*	2	5	8	398	518	525	211	26
*	2	5	10	378	490	1009	350	35
*	2	6	0	264	71	324	0	0
*	2	6	2	299	477	249	169	0
*	2	6	4	276	312	438	60	6
*	2	6	6	430	749	531	159	7
*	2	6	8	196	182	227	194	0
*	2	6	10	693	686	540	251	35
*	2	7	0	267	786	359	0	0
*	2	7	2	398	117	203	2	34
*	2	7	4	223	318	81	114	9
*	2	7	6	376	240	512	163	0
*	2	7	8	224	496	299	55	0
*	2	7	10	336	282	598	148	0
*	2	8	0	287	97	75	180	0
*	2	8	2	483	139	655	357	20
*	2	8	4	385	469	417	73	20
*	2	8	6	238	416	300	232	28
*	2	8	8	325	179	443	337	2
*	2	9	0	267	534	154	180	0
*	2	9	2	458	144	611	27	34
*	2	9	4	549	402	466	102	22
*	2	9	6	620	466	412	347	28
*	2	10	0	827	576	617	0	0
*	2	10	2	330	111	462	221	0
*	2	10	4	211	161	172	116	31
*	4	0	0	1350	438	570	0	0
*	4	0	2	1366	33	500	180	0
*	4	0	4	346	295	677	180	0
*	4	0	6	1103	450	71	180	0
*	4	0	8	1388	808	602	0	0
*	4	0	10	143	89	366	180	0
*	4	0	12	347	364	377	180	0
*	4	1	0	339	277	141	0	0
*	4	1	2	357	237	575	89	0
*	4	1	4	1082	1250	1103	260	4
*	4	1	6	497	763	300	106	39
*	4	1	8	224	1271	675	262	1
*	4	1	10	562	780	742	82	19
*	4	1	12	516	112	495	230	0
*	4	2	0	188	162	538	180	0
*	4	2	2	771	1202	1432	170	0

	H	K	L	|FPH1|	|FPH2|	|FP|	PHASE	
*	4	2	4	516	696	752	226	24
*	4	2	6	657	293	296	130	0
*	4	2	8	642	812	786	201	0
*	4	2	10	383	409	286	178	0
*	4	2	12	474	906	835	56	0
*	4	3	0	704	417	485	0	0
*	4	3	2	1305	698	894	243	14
*	4	3	4	1325	1398	1510	124	17
*	4	3	6	226	278	84	77	0
*	4	3	8	798	542	321	4	0
*	4	3	10	471	194	219	266	2
*	4	4	0	594	1324	477	180	0
*	4	4	2	515	1550	898	314	28
*	4	4	4	487	782	270	325	0
*	4	4	6	625	1137	863	307	27
*	4	4	8	533	532	744	100	13
*	4	4	10	65	250	245	330	2
*	4	5	0	532	564	405	180	0
*	4	5	2	1180	554	670	200	0
*	4	5	4	317	486	452	324	28
*	4	5	6	246	141	86	22	23
*	4	5	8	299	437	268	114	0
*	4	5	10	380	358	135	38	15
*	4	6	0	1195	1113	1269	0	0
*	4	6	2	681	280	599	179	13
*	4	6	4	896	523	869	3	23
*	4	6	6	368	518	339	226	0
*	4	6	8	604	393	472	42	37
*	4	7	0	704	575	682	0	0
*	4	7	2	122	485	255	99	15
*	4	7	4	469	366	263	230	29
*	4	7	6	604	651	681	339	0
*	4	7	8	224	556	208	236	32
*	4	8	0	267	233	536	0	0
*	4	8	2	209	616	343	195	36
*	4	8	4	418	289	293	219	0
*	4	8	6	117	87	350	89	9
*	4	9	0	272	91	139	180	0
*	4	9	2	404	701	553	277	0
*	4	9	4	719	501	630	257	0
*	4	10	0	591	643	387	180	0
*	6	0	0	611	844	979	0	0
*	6	0	2	215	538	445	180	0
*	6	0	4	981	732	689	0	0
*	6	0	6	661	179	594	180	0
*	6	0	8	117	478	186	0	0
*	6	0	10	961	585	1082	180	0
*	6	1	0	358	486	444	180	0

	H	K	L	|FPH1|	|FPH2|	|FP|	PHASE
*	6	1	2	852	729	838	187 0
*	6	1	4	384	251	342	21 0
*	6	1	6	708	497	409	67 11
*	6	1	8	378	467	487	301 16
*	6	1	10	567	338	150	261 37
*	6	2	0	213	314	216	180 0
*	6	2	2	999	162	1302	130 0
*	6	2	4	854	707	813	241 1
*	6	2	6	75	306	262	38 33
*	6	2	8	98	211	109	320 34
*	6	2	10	418	370	478	170 21
*	6	3	0	977	1292	891	0 0
*	6	3	2	517	980	446	216 14
*	6	3	4	296	426	378	189 0
*	6	3	6	378	172	418	305 8
*	6	3	8	203	440	377	36 19
*	6	4	0	577	529	998	180 0
*	6	4	2	370	418	344	23 0
*	6	4	4	312	450	377	133 18
*	6	4	6	304	275	95	54 10
*	6	4	8	160	166	369	70 30
*	6	5	0	847	950	746	0 0
*	6	5	2	1039	1221	960	196 21
*	6	5	4	519	528	399	11 12
*	6	5	6	341	187	246	238 5
*	6	5	8	417	229	281	267 11
*	6	6	0	257	296	153	0 0
*	6	6	2	351	444	578	131 0
*	6	6	4	135	141	254	104 39
*	6	6	6	442	404	360	185 18
*	6	7	0	319	211	467	180 0
*	6	7	2	307	419	385	89 5
*	6	7	4	367	267	470	202 31
*	6	7	6	151	408	217	226 0
*	6	8	0	274	161	344	180 0
*	6	8	2	342	149	463	94 1
*	8	0	0	1783	1548	764	0 0
*	8	0	2	631	329	188	180 0
*	8	0	4	885	374	232	0 0
*	8	0	6	132	201	93	0 0
*	8	1	0	263	399	183	180 0
*	8	1	2	436	821	519	62 38
*	8	1	4	324		148	59 0
*	8	1	6	222	510	594	98 37
*	8	2	0	420	341	310	180 0
*	8	2	2	123	531	158	239 0
*	8	2	4	98	701	399	191 0
*	8	2	6	273	355	470	347 0

	H	K	L	|FPH1|	|FPH2|	|FP|	PHASE	
*	8	3	0	697	100	471	0	0
*	8	3	2	380	533	435	344	0
*	8	3	4	107	93	220	347	0
*	8	4	0	372	1688	982	180	0
*	8	4	2	246	623	453	18	1
*	8	4	4	316	489	534	213	0
*	8	5	0	390	110	146	0	0
*	8	5	2	465	256	239	265	38
*	1	0	3	55	116	156	0	0
*	1	0	5	970	362	219	0	0
*	1	0	7	382	672	289	0	0
*	1	0	9	1163	2112	1467	180	0
*	1	0	11	844	401	894	0	0
*	1	0	13	134	540	690	0	0
*	1	1	3	473	782	372	266	31
*	1	1	5	860	103	638	303	34
*	1	1	7	1930	1557	1662	79	0
*	1	1	9	863	400	563	298	23
*	1	1	11	590	843	890	304	37
*	1	1	13	1239	1446	1276	342	35
*	1	2	1	610	919	646	92	27
*	1	2	3	1033	741	1020	284	18
*	1	2	5	470	301	511	313	0
*	1	2	7	612	664	324	306	34
*	1	2	9	635	778	723	124	0
*	1	2	11	60	441	332	140	0
*	1	2	13	425	610	378	9	9
*	1	3	1	816	692	742	294	0
*	1	3	3	54	173	92	221	0
*	1	3	5	978	1066	1107	137	23
*	1	3	7	257	598	537	326	0
*	1	3	9	51	382	138	351	32
*	1	3	11	603	252	594	150	26
*	1	3	13	535	207	264	284	0
*	1	4	1	1098	1222	1303	261	24
*	1	4	3	1045	1349	1021	335	9
*	1	4	5	1171	942	891	290	37
*	1	4	7	531	484	530	149	0
*	1	4	9	309	601	343	66	6
*	1	4	11	266	132	138	240	11
*	1	5	1	890	838	1045	271	0
*	1	5	3	118	476	504	317	17
*	1	5	5	1791	973	1227	326	31
*	1	5	7	535	222	242	275	0
*	1	5	9	366	571	484	135	11
*	1	5	11	64	304	202	335	16
*	1	6	1	905	423	629	226	26
*	1	6	3	224	583	102	202	26

	H	K	L	|FPH1|	|FPH2|	|FP|	PHASE	
*	1	6	5	610	403	447	285	11
*	1	6	7	471	215	434	54	0
*	1	6	9	396	316	107	102	36
*	1	6	11	228	207	197	265	0
*	1	7	1	621	373	850	115	37
*	1	7	3	750	764	775	266	19
*	1	7	5	409	252	557	144	0
*	1	7	7	632	336	408	98	31
*	1	7	9	469	526	686	245	0
*	1	8	1	918	381	396	274	33
*	1	8	3	99	365	380	283	38
*	1	8	5	715	380	201	213	25
*	1	8	7	924	585	579	48	24
*	1	8	9	257	378	292	85	1
*	1	9	1	795	646	614	272	11
*	1	9	3	1301	347	852	124	0
*	1	9	5	681	136	322	327	0
*	1	9	7	228	353	262	57	8
*	1	10	1	57	168	469	92	0
*	1	10	3	814	189	663	43	39
*	1	10	5	288	344	157	63	0
*	1	11	1	769	264	222	250	34
*	3	0	1	63	257	34	180	0
*	3	0	3	742	740	702	0	0
*	3	0	5	389	155	34	0	0
*	3	0	7	719	1722	1269	180	0
*	3	0	9	373	553	463	180	0
*	3	0	11	358	449	339	0	0
*	3	1	1	363	50	587	288	26
*	3	1	3	435	108	216	142	26
*	3	1	5	1902	1537	1949	210	7
*	3	1	7	851	691	1013	299	0
*	3	1	9	258	248	426	282	20
*	3	1	11	140	353	230	303	32
*	3	2	1	572	413	384	342	19
*	3	2	3	1088	1299	1322	283	36
*	3	2	5	655	855	323	329	22
*	3	2	7	1020	426	598	290	4
*	3	2	9	693	322	558	78	0
*	3	2	11	317	287	528	254	15
*	3	3	1	477	875	881	124	0
*	3	3	3	1040	1117	966	170	0
*	3	3	5	1033	1053	909	285	9
*	3	3	7	759	1156	1189	128	30
*	3	3	9	196	838	657	338	12
*	3	3	11	189	645	354	152	0
*	3	4	1	799	1169	932	291	24
*	3	4	3	1101	1186	1075	128	0

	H	K	L	|FPH1|	|FPH2|	|FP|	PHASE
*	3	4	5	1383	1113	1143	22 16
*	3	4	7	777	508	404	217 27
*	3	4	9	592	535	381	80 33
*	3	4	11	226	216	237	244 30
*	3	5	1	851	631	683	252 22
*	3	5	3	365	407	652	198 22
*	3	5	5	426	60	205	261 32
*	3	5	7	495	420	547	92 0
*	3	5	9	448	179	130	133 0
*	3	5	11	105	480	198	217 2
*	3	6	1	915	804	412	156 4
*	3	6	3	412	111	163	144 7
*	3	6	5	557	358	329	137 5
*	3	6	7	308	250	266	44 21
*	3	6	9	460	329	329	135 0
*	3	7	1	98	389	552	107 27
*	3	7	3	897	668	686	326 25
*	3	7	5	500	435	366	361 0
*	3	7	7	253	452	478	210 22
*	3	7	9	124	632	374	351 18
*	3	8	1	573	379	440	355 16
*	3	8	3	618	894	965	125 0
*	3	8	5	528	651	465	24 0
*	3	8	7	135	132	338	135 38
*	3	9	1	388	739	745	97 14
*	3	9	3	383	139	146	226 0
*	3	9	5	864	321	421	192 0
*	3	10	1	450	334	952	274 38
*	3	10	3	442	468	386	215 31
*	5	0	1	1099	679	795	180 0
*	5	0	3	361	118	139	180 0
*	5	0	5	655	1594	1139	180 0
*	5	0	7	790	254	163	180 0
*	5	0	9	167	420	410	180 0
*	5	0	11	122	374	419	0 0
*	5	1	1	405	286	593	29 13
*	5	1	3	716	1123	569	17 0
*	5	1	5	379	188	255	193 0
*	5	1	7	474	585	361	172 0
*	5	1	9	664	681	838	304 0
*	5	1	11	679	587	615	173 11
*	5	2	1	1071	1037	1072	251 5
*	5	2	3	1153	1037	1356	93 0
*	5	2	5	82	660	204	18 0
*	5	2	7	541	483	820	120 8
*	5	2	9	60	701	565	249 0
*	5	2	11	72	370	320	59 0
*	5	3	1	1229	1177	916	73 6

	H	K	L	|FPH1|	|FPH2|	|FP|	PHASE
*	5	3	3	1823	2145	1544	298 35
*	5	3	5	381	682	497	134 31
*	5	3	7	88	315	241	2 10
*	5	3	9	175	185	168	211 0
*	5	4	1	1160	1413	1288	121 37
*	5	4	3	1551	1625	1312	323 20
*	5	4	5	872	1085	1086	146 15
*	5	4	7	440	667	614	312 25
*	5	4	9	63	342	117	111 0
*	5	5	1	272	379	408	80 5
*	5	5	3	748	838	1001	315 0
*	5	5	5	323	628	821	145 0
*	5	5	7	221	429	347	0 11
*	5	5	9	444	337	122	239 37
*	5	6	1	715	443	348	234 15
*	5	6	3	74	205	424	281 22
*	5	6	5	246	325	140	246 0
*	5	6	7	101	205	378	20 11
*	5	7	1	450	405	429	63 30
*	5	7	3	586	657	562	216 31
*	5	7	5	899	52	305	0 0
*	5	7	7	420	269	303	321 35
*	5	8	1	472	828	542	115 8
*	5	8	3	231	441	204	278 23
*	5	8	5	825	971	578	242 0
*	5	9	1	328	567	274	264 0
*	5	9	3	401	972	428	167 18
*	7	0	1	121	224	460	180 0
*	7	0	3	455	604	265	180 0
*	7	0	5	467	837	895	180 0
*	7	0	7	99	175	454	0 0
*	7	1	1	543	457	556	76 0
*	7	1	3	241	353	267	170 7
*	7	1	5	315	174	314	316 38
*	7	1	7	615	57	223	31 36
*	7	2	1	111	595	274	265 0
*	7	2	3	427	1121	607	53 10
*	7	2	5	350	416	184	68 12
*	7	2	7	383	94	190	152 1
*	7	3	1	696	778	655	235 9
*	7	3	3	827	710	773	34 0
*	7	3	5	343	413	355	163 0
*	7	3	7	353	142	185	310 0
*	7	4	1	354	152	686	291 0
*	7	4	3	181	393	499	156 1
*	7	4	5	725	76	654	327 14
*	7	4	7	448	636	278	307 20
*	7	5	1	494	618	456	156 8

	H	K	L	|FPH1|	|FPH2|	|FP|	PHASE	
*	7	5	3	203	608	390	241	0
*	7	5	5	278	349	321	137	20
*	7	6	1	388	70	379	251	0
*	7	6	3	62	619	571	61	5
*	7	6	5	383	398	650	310	0
*	7	7	1	551	51	391	276	2
*	9	0	1	416	280	376	0	0
*	9	0	3	793	512	462	0	0
*	9	1	1	523	254	222	267	0
*	9	1	3	468	596	467	226	0
*	9	2	1	200	344	89	120	7

APPENDIX IV

Reflection Data to 3.3Å Resolution for the [010] Projection of the $IrCl_6^{3-}$ Derivative

	H	K	L	|FPH|	|FP|		H	K	L	|FPH|	|FP|
*	2	0	1	1160	1027	*	3	0	1	1160	-1036
*	4	0	1	892	-872	*	5	0	1	344	498
*	6	0	1	433	535	*	7	0	1	102	127
*	8	0	1	688	-572	*	9	0	1	867	-734
*	10	0	1	790	766	*	11	0	1	408	-356
*	12	0	1	714	734	*	1	0	2	803	701
*	2	0	2	293	317	*	3	0	2	1797	-1810
*	4	0	2	650	-615	*	5	0	2	1032	1076
*	6	0	2	650	-567	*	7	0	2	446	-415
*	8	0	2	867	-842	*	9	0	2	548	-594
*	10	0	2	1032	-1069	*	11	0	2	459	-356
*	12	0	2	306	-291	*	1	0	3	777	1053
*	2	0	3	879	-536	*	3	0	3	229	355
*	4	0	3	663	749	*	5	0	3	421	-238
*	6	0	3	242	256	*	7	0	3	153	127
*	8	0	3	331	-205	*	9	0	3	905	-853
*	10	0	3	650	-594	*	11	0	3	293	-356
*	12	0	3	306	-335	*	1	0	4	1198	-1083
*	2	0	4	1848	1749	*	3	0	4	969	947
*	4	0	4	1822	1764	*	5	0	4	943	940
*	6	0	4	1134	1085	*	7	0	4	892	-917
*	8	0	4	153	-108	*	9	0	4	408	324
*	10	0	4	115	108	*	11	0	4	115	108
*	12	0	4	102	-108	*	1	0	5	306	-461
*	2	0	5	510	327	*	3	0	5	2218	2050
*	4	0	5	153	152	*	5	0	5	280	-375
*	6	0	5	102	209	*	7	0	5	599	-766
*	8	0	5	650	-615	*	9	0	5	115	-151
*	10	0	5	586	529	*	11	0	5	382	248
*	12	0	5	102	-162	*	1	0	6	930	1161 :
*	2	0	6	408	-569	*	3	0	6	1071	999
*	4	0	6	497	-605	*	5	0	6	446	402
*	6	0	6	293	254	*	7	0	6	306	-246
*	8	0	6	1071	1112	*	9	0	6	1173	1155
*	10	0	6	191	108	*	11	0	6	229	248
*	12	0	6	344	-367	*	1	0	7	828	-771
*	2	0	7	191	-163	*	3	0	7	1568	-1414
*	4	0	7	1708	1721	*	5	0	7	280	338
*	6	0	7	178	90	*	7	0	7	790	857
*	8	0	7	331	356	*	9	0	7	930	939
*	10	0	7	599	507	*	11	0	7	115	108
*	12	0	7	178	-291	*	1	0	8	255	141
*	2	0	8	1657	1525 :	*	3	0	8	803	-777
*	4	0	8	523	619	*	5	0	8	930	-896
*	6	0	8	293	214	*	7	0	8	803	816
*	8	0	8	548	-626	*	9	0	8	280	-378
*	10	0	8	879	853	*	11	0	8	115	108
*	12	0	8	650	745	*	1	0	9	1122	1383

	H	K	L	|FPH|	|FP|		H	K	L	|FPH|	|FP|	
*	2	0	9	1963	-1713	*	3	0	9	421	-299	
*	4	0	9	408	330	*	5	0	9	280	186	:
*	6	0	9	102	177	*	7	0	9	650	691	
*	8	0	9	433	432	*	9	0	9	115	151	
*	10	0	9	306	-270	*	11	0	9	561	-702	
*	12	0	9	102	108	*	1	0	10	1007	-937	
*	2	0	10	306	341	*	3	0	10	739	627	
*	4	0	10	331	254	*	5	0	10	140	-177	
*	6	0	10	497	579	*	7	0	10	370	140	
*	8	0	10	803	885	*	9	0	10	191	-367	
*	10	0	10	574	712	*	11	0	10	268	194	
*	12	0	10	102	108	*	1	0	11	1784	1769	:
*	2	0	11	472	292	*	3	0	11	930	930	:
*	4	0	11	268	258	*	5	0	11	204	263	
*	6	0	11	650	-696	*	7	0	11	650	-691	
*	8	0	11	293	-356	*	9	0	11	166	-248	
*	10	0	11	905	939	*	11	0	11	102	194	
*	12	0	11	752	-831	*	1	0	12	510	-611	
*	2	0	12	1453	1457	*	3	0	12	408	524	
*	4	0	12	854	810	*	5	0	12	663	-704	
*	6	0	12	115	91	*	7	0	12	752	-864	
*	8	0	12	166	108	*	9	0	12	790	-853	
*	10	0	12	828	-788	*	11	0	12	319	194	
*	1	0	13	1478	1601	*	2	0	13	242	215	
*	3	0	13	574	573	*	4	0	13	688	-759	
*	5	0	13	178	-183	*	6	0	13	115	-205	
*	7	0	13	280	216	*	8	0	13	510	615	
*	9	0	13	280	227	*	10	0	13	624	-561	
*	11	0	13	548	464	*	1	0	14	548	529	
*	2	0	14	140	152	*	3	0	14	293	340	
*	4	0	14	319	324	*	5	0	14	535	626	
*	6	0	14	331	389	*	7	0	14	293	356	
*	8	0	14	229	-421	*	9	0	14	841	961	
*	10	0	14	268	270	*	11	0	14	497	389	
*	1	0	15	1529	-1462	*	2	0	15	319	457	
*	3	0	15	153	270	*	4	0	15	1007	-1282	
*	5	0	15	459	583	*	6	0	15	739	-799	
*	7	0	15	459	421	*	8	0	15	166	-108	
*	9	0	15	421	540	*	10	0	15	497	-497	
*	11	0	15	102		*	1	0	16	714	679	
*	2	0	16	408	-466	*	3	0	16	574	-623	
*	4	0	16	408	-405	*	5	0	16	115	-184	
*	6	0	16	153	108	*	7	0	16	421	-486	
*	8	0	16	280	-432	*	9	0	16	382	572	
*	10	0	16	344	-399	*	1	0	17	637	658	
*	2	0	17	1071	1043	*	3	0	17	268	-235	
*	4	0	17	242	-227	*	5	0	17	382	313	
*	6	0	17	255	-345	*	7	0	17	650	756	

	H	K	L	|FPH|	|FP|		H	K	L	|FPH|	|FP|
*	8	0	17	166	194	*	9	0	17	1134	-1252
*	10	0	17	102	-227	*	1	0	18	319	-150
*	2	0	18	421	-193	*	3	0	18	714	531
*	4	0	18	574	-766	*	5	0	18	191	-367
*	6	0	18	344	356	*	7	0	18	574	637
*	8	0	18	319	-432	*	9	0	18	331	248
*	10	0	18	102	-194	*	1	0	19	370	-248
*	2	0	19	714	-874	*	3	0	19	319	-151
*	4	0	19	115	237	*	5	0	19	191	184
*	6	0	19	408	-399	*	7	0	19	115	108
*	8	0	19	663	702	*	9	0	19	918	1112
*	10	0	19	382	432	*	1	0	20	293	-302
*	2	0	20	331	302	*	3	0	20	1249	-1295
*	4	0	20	484	-356	*	5	0	20	191	108
*	6	0	20	229	108	*	7	0	20	472	432
*	8	0	20	930	1101	*	9	0	20	497	-486
*	10	0	20	89		*	1	0	21	574	637
*	2	0	21	701	-540	*	3	0	21	510	-345
*	4	0	21	714	864	*	5	0	21	1096	-1101
*	6	0	21	918	-1004	*	7	0	21	408	-583
*	8	0	21	191		*	9	0	21	268	399 :
*	1	0	22	956	982	*	2	0	22	115	108
*	3	0	22	319	-291	*	4	0	22	115	216
*	5	0	22	624	540	*	6	0	22	472	529
*	7	0	22	599		*	8	0	22	280	
*	9	0	22	803	853 :	*	1	0	23	3352	-3627
*	2	0	23	956	-950	*	3	0	23	1313	1263
*	4	0	23	115	108	*	5	0	23	293	335
*	6	0	23	459	-561	*	7	0	23	102	
*	8	0	23	675		*	1	0	24	930	-928
*	2	0	24	459	367	*	3	0	24	1835	-1997
*	4	0	24	421	594	*	5	0	24	280	248
*	6	0	24	153		*	7	0	24	229	
*	8	0	24	331		*	1	0	25	599	583
*	2	0	25	1466	-1447	*	3	0	25	191	108
*	4	0	25	650	-842	*	5	0	25	408	367
*	6	0	25	268		*	7	0	25	484	.
*	1	0	26	344	-108	*	2	0	26	1504	1533
*	3	0	26	1185	1317	*	4	0	26	382	
*	5	0	26	153		*	6	0	26	204	
*	7	0	26	89		*	1	0	27	217	108
*	2	0	27	293	-108	*	3	0	27	688	-810
*	4	0	27	242		*	5	0	27	280	
*	6	0	27	382		*	1	0	28	217	-108
*	2	0	28	943	982	*	3	0	28	331	335
*	4	0	28	102		*	5	0	28	102	
*	1	0	29	102	-162	*	2	0	29	637	
*	3	0	29	268		*	4	0	29	140	

	H	K	L	|FPH|	|FP|		H	K	L	|FPH|	|FP|
*	1	0	30	319		*	2	0	30	306	
*	3	0	30	675		*	2	0	0	229	131
*	3	0	0	561	441 :	*	4	0	0	472	-478
*	5	0	0	446	371	*	6	0	0	166	-173
*	7	0	0	382	-473	*	8	0	0	191	-97
*	9	0	0	586	594	*	10	0	0	535	443
*	11	0	0	408	-421	*	12	0	0	637	637
*	2	0	-1	663	-718	*	3	0	-1	1886	1888 :
*	4	0	-1	1453	1463	*	5	0	-1	331	265
*	6	0	-1	472	430	*	7	0	-1	102	127
*	8	0	-1	191	-97	*	9	0	-1	191	151
*	10	0	-1	1198	1166	*	11	0	-1	255	-227
*	12	0	-1	306	270	*	1	0	-2	892	-938
*	2	0	-2	64	116	*	3	0	-2	867	-935
*	4	0	-2	408	440 :	*	5	0	-2	650	772
*	6	0	-2	905	690	*	7	0	-2	548	560
*	8	0	-2	663	605	*	9	0	-2	344	-108
*	10	0	-2	497	410	*	11	0	-2	331	270
*	12	0	-2	854	-885	*	0	0	-3	905	
*	1	0	-3	293	-59	*	2	0	-3	484	604
*	3	0	-3	102	302	*	4	0	-3	1759	-1590
*	5	0	-3	739	798	*	6	0	-3	1593	1545
*	7	0	-3	255	213	*	8	0	-3	599	712
*	9	0	-3	166	108	*	10	0	-3	1389	-1328
*	11	0	-3	115	162	*	12	0	-3	548	-572
*	1	0	-4	395	260	*	2	0	-4	1173	1270
*	3	0	-4	1071	-1056	*	4	0	-4	421	-206
*	5	0	-4	421	418	*	6	0	-4	382	247
*	7	0	-4	421	-330	*	8	0	-4	816	896
*	9	0	-4	344	335	*	10	0	-4	446	-313
*	11	0	-4	854	810	*	12	0	-4	484	-389
*	1	0	-5	586	-535	*	2	0	-5	905	-813
*	3	0	-5	1007	1074	*	4	0	-5	1007	-1029
*	5	0	-5	191	-115	*	6	0	-5	790	-725
*	7	0	-5	994	1103	*	8	0	-5	586	-551
*	9	0	-5	191	-151	*	10	0	-5	523	-464
*	11	0	-5	191	-270	*	12	0	-5	675	637
*	1	0	-6	395	-608	*	2	0	-6	918	634
*	3	0	-6	268	-365	*	4	0	-6	370	166
*	5	0	-6	663	754	*	6	0	-6	255	-139
*	7	0	-6	217	91	*	8	0	-6	153	-151
*	9	0	-6	433	-497	*	10	0	-6	1338	-1295
*	11	0	-6	268	-313	*	12	0	-6	293	-313
*	1	0	-7	370	322	*	2	0	-7	76	72
*	3	0	-7	153	160	*	4	0	-7	714	648
*	5	0	-7	217	-195	*	6	0	-7	153	205
*	7	0	-7	344	250	*	8	0	-7	1325	-1328
*	9	0	-7	459	475	*	10	0	-7	280	227

	H	K	L	|FPH|	|FP|		H	K	L	|FPH|	|FP|
*	11	0	-7	382	-389	*	12	0	-7	229	227
*	1	0	-8	650	772	*	2	0	-8	1606	1550
*	3	0	-8	344	281	*	4	0	-8	663	-763
*	5	0	-8	535	-583	*	6	0	-8	153	-127
*	7	0	-8	816	866	*	8	0	-8	357	281
*	9	0	-8	790	-907	*	10	0	-8	229	108
*	11	0	-8	115	162	*	12	0	-8	446	-497
*	0	0	-9	2281	-2324	*	1	0	-9	115	395
*	2	0	-9	777	-801	*	3	0	-9	319	-353
*	4	0	-9	943	-884	*	5	0	-9	828	-743
*	6	0	-9	153	92	*	7	0	-9	115	-140
*	8	0	-9	115	-216	*	9	0	-9	115	-248
*	10	0	-9	115	-108	*	11	0	-9	115	162
*	12	0	-9	548	-540	*	1	0	-10	2103	-2137
*	2	0	-10	497	465	*	3	0	-10	599	580
*	4	0	-10	344	-256	*	5	0	-10	816	-785
*	6	0	-10	624	-696	*	7	0	-10	217	-291
*	8	0	-10	433	475	*	9	0	-10	255	270
*	10	0	-10	191	-108	*	11	0	-10	408	-399
*	12	0	-10	102	108	*	1	0	-11	790	-767
*	2	0	-11	1071	1065	*	3	0	-11	331	-350
*	4	0	-11	765	694	*	5	0	-11	739	798
*	6	0	-11	561	734	*	7	0	-11	115	281
*	8	0	-11	561	648	*	9	0	-11	1109	-1231
*	10	0	-11	688	777	*	11	0	-11	523	-399
*	12	0	-11	89	108	*	0	0	-12	535	468
*	1	0	-12	510	317	*	2	0	-12	1071	900
*	3	0	-12	726	586	*	4	0	-12	306	-125
*	5	0	-12	306	223	*	6	0	-12	331	-434
*	7	0	-12	115	-216	*	8	0	-12	229	-291
*	9	0	-12	688	605	*	10	0	-12	370	194
*	11	0	-12	688	-723	*	1	0	-13	981	875
*	2	0	-13	574	-716	*	3	0	-13	421	233
*	4	0	-13	472	-541	*	5	0	-13	217	123
*	6	0	-13	191	-259	*	7	0	-13	115	151
*	8	0	-13	650	-745	*	9	0	-13	1644	1652
*	10	0	-13	599	529	*	11	0	-13	510	-680
*	1	0	-14	675	614	*	2	0	-14	752	820
*	3	0	-14	510	-522	*	4	0	-14	102	-143
*	5	0	-14	115	-154	*	6	0	-14	879	-972
*	7	0	-14	344	-529	*	8	0	-14	191	108
*	9	0	-14	395	421	*	10	0	-14	268	-162
*	11	0	-14	561	486	*	0	0	-15	1058	-944
*	1	0	-15	930	-771	*	2	0	-15	140	293
*	3	0	-15	421	556	*	4	0	-15	102	130
*	5	0	-15	446	594	*	6	0	-15	739	788
*	7	0	-15	115	108	*	8	0	-15	395	-216
*	9	0	-15	153	-108	*	10	0	-15	561	-669

	H	K	L	|FPH|	|FP|		H	K	L	|FPH|	|FP|
*	11	0	-15	586	561	*	1	0	-16	331	434
*	2	0	-16	535	-481	*	3	0	-16	956	-898
*	4	0	-16	446	493	*	5	0	-16	153	108
*	6	0	-16	166	-108	*	7	0	-16	816	-853
*	8	0	-16	255	227	*	9	0	-16	370	335
*	10	0	-16	790	-810	*	1	0	-17	421	-566
*	2	0	-17	102	108	*	3	0	-17	153	91
*	4	0	-17	153	-313	*	5	0	-17	459	410
*	6	0	-17	867	918	*	7	0	-17	191	248
*	8	0	-17	892	-950	*	9	0	-17	344	-270
*	10	0	-17	102	108	*	0	0	-18	854	716
*	1	0	-18	242	152	*	2	0	-18	854	-901
*	3	0	-18	115	-108	*	4	0	-18	153	-108
*	5	0	-18	370	-356	*	6	0	-18	612	-615
*	7	0	-18	472	540	*	8	0	-18	191	227
*	9	0	-18	306	248	*	10	0	-18	765	605
*	1	0	-19	446	-507	*	2	0	-19	816	756
*	3	0	-19	115	108	*	4	0	-19	293	-356
*	5	0	-19	1593	1619	*	6	0	-19	739	-691
*	7	0	-19	574	-648	*	8	0	-19	612	712
*	9	0	-19	306	-389	*	10	0	-19	268	
*	1	0	-20	255	-302	*	2	0	-20	446	-410
*	3	0	-20	229	108	*	4	0	-20	535	-529
*	5	0	-20	395	-291	*	6	0	-20	1338	1166
*	7	0	-20	255	-248	*	8	0	-20	841	-831
*	9	0	-20	229	162	*	10	0	-20	89	
*	0	0	-21	395	-551	*	1	0	-21	191	184
*	2	0	-21	229	108	*	3	0	-21	306	-108
*	4	0	-21	969	-982	*	5	0	-21	1134	-1101
*	6	0	-21	1083	1058	*	7	0	-21	191	-108
*	8	0	-21	191	-108	*	9	0	-21	217	-108
*	1	0	-22	930	-1187	*	2	0	-22	166	151
*	3	0	-22	510	518	*	4	0	-22	943	950
*	5	0	-22	1695	1641	*	6	0	-22	421	-432
*	7	0	-22	1389	1403	*	8	0	-22	204	-162
*	9	0	-22	89	108	*	1	0	-23	1313	-1457
*	2	0	-23	752	-928	*	3	0	-23	472	-432
*	4	0	-23	344	-313	*	5	0	-23	688	712
*	6	0	-23	153	194	*	7	0	-23	714	756
*	8	0	-23	574	-680	*	1	0	-24	306	313
*	2	0	-24	1045	-961	*	3	0	-24	433	-583
*	4	0	-24	191	108	*	5	0	-24	675	-637
*	6	0	-24	459	-486	*	7	0	-24	382	356
*	8	0	-24	586	-712	*	1	0	-25	1861	-1652
*	4	0	-25	115	-108	*	5	0	-25	395	194
*	2	0	-25	1274	-1295	*	3	0	-25	191	-313
*	6	0	-25	178		*	7	0	-25	599	
*	1	0	-26	293	-108	*	2	0	-26	790	626

	H	K	L	|FPH|	|FP|		H	K	L	|FPH|	|FP|
*	3	0	-26	446	-399	*	4	0	-26	574	-669
*	5	0	-26	178	-108	*	6	0	-26	268	
*	7	0	-26	548		*	0	0	-27	421	-270
*	1	0	-27	331	108	*	2	0	-27	1211	1263
*	3	0	-27	446	-443	*	4	0	-27	472	421
*	5	0	-27	548	-659	*	6	0	-27	229	
*	1	0	-28	726	723	*	2	0	-28	217	
*	3	0	-28	523	443	*	4	0	-28	777	-810
*	5	0	-28	102		*	1	0	-29	229	367
*	2	0	-29	574	-572	*	3	0	-29	204	
*	4	0	-29	370		*	0	0	-30	1007	-1069
*	1	0	-30	204	-270	*	2	0	-30	637	
*	3	0	-30	637							

NO PHASE INFORMATION IS AVAILABLE FOR THOSE MARKED (:).

REFERENCES

Anderson, P.J., Gibbons, I. and Perham, R.N. (1969) Eur.J.Biochem. 11 503.

Anderson, P.J., and Perham, R.N. (1970) Biochem. J. 117, 291.

Armstrong, J.M., McKenzie, H.A. & Sawyer, W.H. (1967). Biochim. Biophys. Acta. 147,60.

Arndt, U.W. & Willis, B.T.M. (1966). Single Crystal Diffractometry Cambridge University Press.

Ascaffenburg, R. & Drewry, J. (1955). Nature 176, 218.

Aschaffenburg, R.E. & Drewry, J. (1957a). Biochem. J. 65, 273.

Aschaffenburg, R. & Drewry, (1957b) Nature 180, 376.

Aschaffenburg, R. Green, D.W. & Simmons, R.M. (1965). J.Mol.Biol. 13, 194.

Askonas, B.A. (1957). Biochem. J. 48, 42.

Babkov, A.V. (1967). Dok.Akad.Nauk.SSSR. 177, 337 from C.A. 68 56032h.

Baranowski, T. (1939). Z.Physiol.Chem. 260, 43.

Bardet, L. & Bontoux, J. (1961). Trav.Soc.Pharm. Montpelier 21, 111.

Basch, J.J. & Timasheff, J.N. (1967). Arch.Biochem.Biophys. 118, 37.

Basolo, F. & Pearson, R.G. (1967). Mechanisms of Inorganic Reactions (2nd Ed.) Wiley, New York.

Bealing, J.F. Czok, R., Eckert, L. & Jäger, I. (1960). Unpublished per Czok, R. & Bucher, T. (1961).

Beck. W.S. (1954). J.Biol.Chem. 212, 847.

Beisenherz, G., Boltze, H.J., Bucher, T.,Czok, R., Garbade, K.H., Meyer-Arendt, E. & Pfleeiderer, G. (1953). Z.Naturforsch. 8b 555.

Bell, K. (1962). Nature 195, 705.

Bell, K. & McKenzie, H.A. (1964). Nature 204, 1275.

Bell, K. & McKenzie, H.A. (1967). Biochem.Biophys.Acta. 147, 109

Benesch, R.E., Lardy, H.A. & Benesch, R. (1955). J.Biol.Chem. 216, 663.

Bergmeyer, H.U. (1965). Methods of Enzymatic Analysis. Academic Press, New York.

Bernfield, P., Berkeley, B.J. & Bieber, R.E. (1965). Arch.Biochim. Biophys. 111, 38.

Bertinotti, C. & Bertinotti, A. (1968). Compt.Rend.Acad.Sci.Paris t267B, 1227.

Bertinotti, C. & Bertinotti, A. (1970). Acta.Cryst. B26, 422.

Bijvoet, J.M. (1954). Nature 173, 888.

Bishop, W.H. & Richards, F.M. (1968). J.Mol.Biol. 38, 315.

Blake, C.C.F. (1968). Adv. in Protein Chemistry. 23, 59.

Blake, C.C.F., Koenig, D.F., Mair, G.A., North A.C.T., Phillips, D.C., & Sarma, V.R. Nature 206, 757. (1965).

Blow, D.M. (1958). Proc.Roy.Soc. A. 247, 302.

Blow, D.M. & Crick, F.H.C. (1959). Acta.Cryst. 12, 794.

Blow, D.M., Rossman, M.G. & Jeffrey, B.A. (1964). J.Mol.Biol. 8, 65.

Blow, D.M. & Steitz, T.A. (1970). Ann.Rev.Biochem. 39, 63.

Bokhoven, C., Schoone, J.C. & Bijvoet, J.M. (1951). Acta. Cryst. 4, 275.

Boyes-Watson, J., Davidson, E. & Perutz, M.F. (1947). Proc.Roy.Soc. A191, 83.

Brignon, G., Ribadeau-Dumas, B., Garnier, J., Pantaloni, D., Guinand, S., Basch, J.J. & Timasheff, S.N. Arch.Biochem.Biophys. 129, 720. (1969).

Britton, H.T.S. & Dodd, E.N. (1935). J.Chem.Soc. 1935, 100.

Browne, W.J., North, A.C.T., Phillips, D.C., Brew, K., Vanaman, T.C., & Hill, R.L. J.Mol.Biol. 42, 65.

Bruns, F. (1954). Biochem. Z. 325, 156.

Bruns, F.H. & Bergmeyer, H.U. (1965). Methods in Enzymatic Analysis p724. Ed: Bergmeyer, H.U. Academic Press, New York.

Buerskens, P.T., Cras, J.A. & Steggarda, J.J. (1968). Inorg. Chem. 7, 810.

Bull, H.B. (1946). J.Am.Chem.Soc. 68, 746.

Bull, H.B. & Currie, B.T. (1946) J.Am.Chem. Soc. 68, 742.

Calvin, M. (1954). Glutathione: a Symposium. Colowick et al. (Eds.) Academic Press.

Carlisle, C.H. & Palmer, R.A. (1962). Acta.Cryst. 15, 129.

Castellino, F.J. & Barker, R. (1966). Biochem.Biophys.Res.Commun. 23, 182.

Cattalini, L., Marangoni, G & Martelli, M. (1968). Inorg.Chem. 7, 1145.

Cattalini, L., Orio, A. & Tobe, M.L. (1967). J.Am.Chem.Soc. 89 3130.

Cecil, R. & McPhee, J.R. (1959). Adv.Prot.Chem. 14 256.

Cecil, R. & Ogston, A.G. (1949). Biochem. J. 44, 33.

Chan, W., Morse, D.E. & Horecker, B.L. (1967). Proc.Natl.Acad.Sci.U.S. (Wash.) 57, 1013.

Christen, P. Rensing, U., Schmid, A. & Leuthardt (1966). Helv. Chim.Acta. 49, 1872.

Cohn, E.J. (1925). Physiol. Revs. 5, 349.

Cohn, E.J., Strong, L.E., Hughes, W.L., Mulford, D.J., Ashwood, J.N. Melin, M. & Taylor, H.L. (1946). J.Am.Chem.Soc. 68, 459.

Crestfield, A.M., Moore, S. & Stein, W.H. (1963). J.Biol.Chem. 238, 622.

Crick, F.H.C. & Magdoff, B.S. (1956). Acta.Cryst. 9, 901.

Cullis, A.F., Muirhead, H., Perutz, M.F., Rossman, M.G. & North,A.C.T. (1962). Proc.Roy.Soc. A265, 15.

Cunningham, L.W. & Neunke, B.J. (1959). J.Biol.Chem. 234, 1447.

Czok, R. & Bücher, T. (1960). Adv.in Prot. Chem. 15, 315.

Davidson, B. & Fasman, G.D. (1967). Biochemistry, 6 1630.

Davidson, B., Tooney, N. & Fasman, G.D. (1966). Biochem.Biophys. Res.Commun. 23, 156.

Davidson, C.M. & Jameson, R.F. (1965). Trans.Farad.Soc. 61 2462.

Dawson, R.M.C., Elliott, D.C., Elliott, W.H. & Jones, K.M. (1959) Data for Biochemical Research. Oxford, Clarendon Press,1st Edn.

Deal, W.C., Rutter, W.J. & van Holde, K.E. (1963). Biochemistry, 2 246.

Dickerson, R.E., Eisenberg, D., Varnum, J. & Kopka, M.L. (1969). J.Mol.Biol. 45, 77.

Dickerson, R.E. & Geis, I. (1969). The Structure and Action of Proteins. Harper & Row, New York.

Dickerson, R.E.,Kendrew, J.G. & Strandberg, B.E. (1961). Acta. Cryst., 14. 1188.

Dickerson, R.E., Kopka, M.L., Borders, C.L., Varnum, J., Weinzierl, J.E., & Margoliash, I. (1967). J.Mol.Biol. 29, 77.

Dickerson, R.E., Kopka, M.L., Varnum, J.C., & Weinzierl, J.E. (1967). Acta Cryst. 23, 511

Dickerson, R.E., Weinzierl, J.E. & Palmer, R.A. (1968). Acta Cryst. B24,997.

Disteche, A. (1948). Biochim.Biophys.Acta 2, 265.

Dixon, M. & Webb, E.C. (1961). Adv. Prot.Chem. 16, 197.

Dixon, M. & Webb, E.C. (1964). Enzymes, 2nd Edition, Longmans, London.

Dounce, A.L., Barnett, S.R. & Beyer, G.T. (1950). J.Biol. Chem. 185, 769.

Dreschler, E.R., Kowalsky, A.G. & Boyer, P.D. (1959). J.Biol. Chem. 234, 2627.

Drenth, J., Kloosterman, D., v.d. Woude, J., Croon, H.C., & v.Zwet, L.C.M. (1965). J.Sci.Instr. 42, 222.

Dunnhill, P. (1963). Ph.D. Thesis presented to London University.

Dunnhill, P. & Green, D.W. (1965). J.Mol.Biol. 15, 147.

Dunnhill, P., Green, D.W. & Simmons, R.M. (1966). J.Mol.Biol. 22, 135.

Durrant, P.J. & Durrant, B. (1962). Introduction to Advanced Inorganic Chemistry. Longmans, London.

Eagles, P.A.M., Johnson, L.N., Joynson, M.A. McMurray, C.H. & Gutfreund, H. (1969). J.Mol.Biol. 45, 533.

Edelstein, S.J. & Schachman, H.K. (1966). Fed.Proc. 25, 412.

El-Awady, A.D., Bounsall, E.J. & Garner, C.S. (1967). Inorg.Chem. 6, 79.

Eldjarn, L. & Pihl, A. (1957). J.Am.Chem.Soc. 79, 4589.

Ellman, G.L. (1959). Arch.Biochem.Biophys. 82, 70.

Enzyme Nomenclature. (1965). Recommendations (1964) of the International Union of Biochemistry, Elsevier, Amsterdam.

Foxwell, C.J., Cran, E.J., Baron, D.N. (1966). Biochem.J. 100, 44p.

Frank, G. & Braunitzer, G. (1967). Hoppe Seyler Z.f.Phys.Chem. 348,1691.

Frank, G. & Braunitzer, G. (1968). Hoppe Seyler Z.f.Phys.Chem. 349, 1456.

Freeman, H.C. (1967). Adv.Prot.Chem 22, 257.

Gibbons, I., Anderson, P.J., Perham, R.N. (1970). FEBS Letters. 10, 49.

Ginsburg, A. (1966). Arch.Biochem.Biophys. 117, 445.

Ginsburg, A. & Mehler, A.H. (1966). Biochemistry 5, 2623.

Ginstrup, O & Leden I. (1968). Acta.Chem.Scand. 22, 1163.

Gordon, W.G., Basch, J.J. & Kalan, E.B. (1961). J.Biol.Chem. 236 2908.

Gorin, G. & Doughty, G. (1968) Arch.Biochem.Biophys. 126, 547.

Grazi, E. Cheng, T., Horecker, B.L. (1962a). Biochem.Biophys.Res. Commun. 7, 250.

Grazi, E., Rowley, P.T., Cheng, T., Tchola, O.& Horecker, B.L. (1962b) Biochem.Biophys.Res.Commun. 9, 38.

Green, A.A. (1931). J.Biol.Chem. 93,495, 517.

Green, A.A. (1932). J.Biol.Chem. 95,47.

Green, D.W., Camerman, A. Komorowski, E.S. & Simmons, R.M. (1971). To be published.

Green, D.W., Ingram, V.M. & Perutz, M.F. (1954). Proc.Roy.Soc. A225, 287.

Green, D.W., North, A.C.T. & Ashaffenburg, R. (1956). Biochim.Biophys. Acta. 21, 583.

Greenfield, N., Davidson, E. & Fasman, G.D. (1967). Biochemistry, 6, 1630.

Greenfield, N. & Fasman, G.D. (1969). Biochemistry, 8, 4108

Groves, M.L., Hipp, N.J. & McMeekin, T.L. (1957). J.Am.Chem.Soc. 73, 1957.

Guertler, B. & Leuthardt, F. (1970). Helv.Chim.Acta. 53, 654.

Halwer, M. Nutting, G.C. & Brice, B.A. (1951). J.Am.Chem.Soc. 73, 2786.

Hamilton, W.C., Rollet, J.S., & Sparks, R.A. (1965). Acta Cryst. 18, 129.

Hardy, P.M. (1969). Amino acids, Peptides and Proteins Vol. I. Specialist Periodical Reports; The Chemical Society, London.

Harker, D. (1953). Acta Cryst. 6, 231.

Hart, R.G. (1961). Acta Cryst. 14, 1194.

Hartman, F.C. & Barker, R. (1965). Biochemistry 4, 1068.

Hass, L.F. (1964). Biochemistry 3, 535.

Haurowitz, F. (1950). Chemistry & Biology of Proteins, p.91. New York: Academic Press.

Herbert, D., Gordon, H., Subrahmanyan, V. & Green, D.E. (1940). Biochem.J. 34, 1109.

Herskovits, J.J., Masters, C.J., Wasserman, P.M. & Kaplan, N.O. (1967). Biochem.Biophys.Res.Commun. 26, 24.

Hoffee, P., Lai, C.Y., Pugh, E.L. & Horecker, B.L. (1967). Arch. Biochem.Biophys., 122, 196.

Hofmeister, F. (1887). Arch.f.Exptl.Path.Pharm. 24, 247.

Holmes, K.C. and Blow, D.M. (1966). Methods in Biochemical Analysis, 13, 113 and Interscience: The use of Xray Diffraction in the Study of Protein and Nucleic Acid Structure.

Horecker, B.L. (1966). Lecture given to the Molecular Enzymology Group of the Biochemical Society, London, 9th November, 1966.

James, R.W. (1954). The Optical Principles of the Diffraction of X-rays. Bell, London.

Jencks, W.P. (1963). Methods of Enzymology Vol. VI p914 Eds. Colowick, S.P. & Kaplan, N.O. Academic Press.

Jirgensons, C.K. (1961). Arch.Biochim.Biophys. 92, 216.

Jocelyn, P.C., (1967). Eur.J.Biochem. 2, 327.

Kalan, E.B., Greenberg, R. & Walter, M. (1965). Biochemistry, 4, 991.

Kartha, G. & Parthasarathy, R. (1965). Acta Cryst. 18, 745,749.

Kawahara, K. & Tanford, C. (1966). Biochemistry, 5, 1578.

Kendrew, J.C. (1962). Enzyme Models & Enzyme Structures, Brookhaven Symposia in Biology No. 15.

King, M.V. (1954). Acta Cryst. 7, 601.

Klotz, I.M., Langerman, N.R. & Darnall, D.W. (1970). Ann Rev. Biochem. 39, 25.

Kobashi, K., Lai, C.Y. & Horecker, B.L. (1966). Arch.Biochem. Biophys. 117, 437.

Kochman, M.,Penhoet, E. & Rutter, W.J. (1968). Fed.Proc. 17, 590.

Koida, M., Lai, C.Y. & Horecker, B.L. (1969). Arch.Biochim. Biophys. 134, 623.

Koida, M. & Lai, C.Y. (1969). Fed.Proc. 28, 3416.

Komorowski, E.S. (1971). Ph.D. Thesis submitted to Edinburgh University.

Koshland, D.E. (1963). Science, 142, 1533.

Kowal, J., Cremona, T. & Horecker, B.L. (1965). J.Biol.Chem. 240, 2485.

Kowalsky, A.G. & Boyer, P.D. (1960). J.Biol.Chem. 235, 604.

Kraut, J., Sieker, L.S., High, D.F. & Freer, S.J. (1962). Proc. Natl.Acad.Sci., Wash. 48, 1417.

Lai, C.Y. (1968). Arch.Biochem.Biophys. 128, 202.

Lai, C.Y. & Chen, C. (1968). Arch.Biochem.Biophys. 128, 212.

Lai, C.Y., Chen, C. & Horecker, B.L. (1970). Biochem.Biophys.Res. Commun. 40, 461.

Lai, C.Y., Hoffee, P. & Horecker, B.L. (1967). Methods in Enzymology. XI 667. Ed. Hirs. C.H.W. Academic Press, New York.

Lai, C.Y., & Hoffee, P. (1966). Fed.Proc. 25, 408.

Lai, C.Y. & Horecker, B.L. (1968). Personal communication.

Lai, C.Y., Martinez de Dretz, G., Bacila, M., Marinello, E. & Horecker, B.L. (1968). Biochem.Biophys.Res.Commun. 30, 665.

Lai, C.Y., Tchola, O., Cheng, T. & Horecker, B.L. (1965). J.Biol.Chem. 240, 1347.

Latimer, W.H. (1952). Oxidation Potentials. 2nd Ed. Prentice-Hall, New York.

Lebherz, H.G. & Rutter, W.J. (1969). Biochemistry, 8, 109.

Linderstrøm-Lang, K.V. & Jacobsen, C.F. (1941). J.Biol.Chem. 137, 443.

Lipscomb, W.N., Coppola, J.C., Hartsuck, J.A., Ludwig, M.L., Muirhead, H., Searl, J. & Steitz, T.A. (1966). J.Mol.Biol. 19, 423.

Low, B.W. & Richards, F.M. (1952). J.Am.Chem.Soc. 74, 1660.

Lundberg, B. (1965). Acta Cryst. 18, 576.

Luzzati, V. (1955). Acta Cryst. 8, 795.

McKenzie, H.A. (1967). Adv. in Prot.Chem. 22, 55.

McKenzie, H.A. & Sawyer, W.H. (1967). Nature, 214, 1101.

McKenzie, H.A., Sawyer, W.H. & Smith, M.B. (1967). Biochim. Biophys.Acta. 147, 73.

Magar, M.E. (1968). Biochemistry 7, 617.

Matthews, B.W. (1966a). Acta Cryst. 20, 82.

Matthews, B.W. (1966b). Acta Cryst. 20, 230.

Matthews, B.W. (1968). J.Mol.Biol. 33, 491.

Matthews, B.W., Sigler, P.B., Henderson, R. & Blow, D.M. (1967). Nature, 214, 652.

Mehler, A.H. (1963). J.Biol.Chem. 238, 100.

Mehler, A.H. & Cusic, M.E. (1967). Science. 155, 1011.

Meyerhof, O. (1957). per Methods in Enzymatic Analysis, Ed. Bergmeyer, H.-U., p246. Academic Press, New York.

Meyerhof, O. & Lohmann, K. (1934a). Biochem. Z. 273, 413.

Meyerhof, O. & Lohmann, K. (1934b). Naturwissenschaften, 22, 220.

Moews, P.C. & Bunn, C.W. (1970). J.Mol.Biol. 54, 395.

Moffitt, W. & Yang, J.T. (1956). Proc.Natl.Acad.Sci.U.S.(Wash.). 42, 596.

Moore, W.J. (1963). Physical Chemistry, Fourth Edition. Longmans, London.

Morse, D.E., Chan, W. & Horecker, B.L. (1967). Proc.Natl.Acad.Sci. (Wash.) 58, 628.

Morse, D.E. & Horecker, B.L. (1968). Adv. in Enzymology. 31, 125.

Nicholas, P.C., Bachelard, H.S. (1969). Biochem.J. 112, 587.

North, A.C.T. (1959). Acta Cryst. 12, 512.

North, A.C.T. (1964). J.Sci.Instr. 41, 42.

North, A.C.T. (1965). Acta Cryst. 18, 212.

North, A.C.T. & Phillips, D.C. (1969). Prog. in Biophysics and Mol.Biol. 19, Part I 5.

North, A.C.T., Phillips, D.C. & Matthews, F.S. (1968). Acta Cryst. A24, 351.

Nozaki, Y., Bunville, L.G. & Tanford, C. (1959). J.Am.Chem.Soc. 81, 5523.

Palmer, A.H. (1934). J.Biol.Chem. 104, 359.

Pantaloni, D. (1965). Doctoral Thesis presented to University of Paris.

Parthasarathy, S. & Ramachandran, G.N. (1966). Acta. Cryst. 21, 163.

Parumareddi, J.R., Liehr, A.D. & Adamson, A.W. (1963). J.Am.Chem.Soc. 85, 249

Pauling, L., Corey, R.B. & Branson, H.R. (1957). Proc.Natl.Acad.Sci. U.S. 37, 205.

Pauling, L. & Corey, R.B. Proc.Natl.Acad.Sci.U.S. 37, 729. (1957).

Penhoet, E., Kochman, M. & Rutter, W.J. (1969). Biochemistry. 8, 4391, 4396.

Penhoet, E., Kochman, M., Valentine, R. & Rutter, W.J. (1967). Biochemistry 6, 2940.

Penhoet, E., Rajkumar, T. & Rutter, W.J. (1966). Proc.Natl.Acad. Sci.U.S. 56, 1275.

Perutz, M.F. (1946). Disc. Farad. Soc. B42, 187.

Perutz, M.F. (1956). Acta Cryst. 13, 221.

Perutz, M.F. (1968). J.Crystal Crowth. 2, 54.

Perutz, M.F. & Lehmann, H. (1968). Nature 219, 902.

Phillips, D.C. (1966a). Advances in Structure Determination by Diffraction Methods. Ed. R.Brill & R.Mason. Vol. II. p.75.

Phillips, D.C. (1966b). Scientific American 215 (November) 78.

Phillips, R.F. & Powell, H.M. (1939). Proc.Roy.Soc. A173, 147.

Piez, K.A., Davie, E.W., Folk, J.E. & Gladner, J.A. (1961). J.Biol. Chem. 236, 2912.

Racker, E. (1947). J.Biol.Chem. 167, 843.

Ramachandran, G.N. (1964). Advanced Methods in Crystallography, Academic Press, London.

Remick, A.E. (1947). J.Am.Chem.Soc. 69, 94.

Richards, O.C. & Rutter, W.J. (1961). J.Biol.Chem. 236, 3185.

Rose, I.A., O'Connell, E.L. & Mehler, A.H. (1965). J.Biol.Chem. 240, 1758.

Rossi, G.L. & Bernhard, S.A. (1970). J.Mol.Biol. 49, 85.

Rossman, M.G. (1960). Acta Cryst. 13, 221.

Rossman, M.G. (1961). Acta Cryst. 14, 383.

Rutter, W.J., Richards, O.C. & Woodfin, B.M. (1961). J.Biol.Chem. 236, 3193.

Sarkar, P.K. & Doty, P. (1966). Proc.Natl.Acad.Sci. 55, 981.

Schellman, J.A. (1958). Compt.Rend.Trav.Lab.Carls.Ser.Chim. 30, 395.

Schmir, G.L. (1965). J.Am.Chem.Soc. 87, 2743.

Schroeder, W.A. (1968). The Primary Structure of Proteins. Harper & Row, New York.

Scott, E.M., Duncan, I.W. & Ekstrand, V. (1963). J.Biol.Chem. 238, 3928.

Scouloudi, H. (1969). J.Mol.Biol. 40, 353.

Schachman, H.K. & Edelstein, S.J. (1966). Biochemistry 5, 2681.

Sheldrick, B. (1970). Quart.Rev. 24, 454.

Shimizu, H. & Ozawa, H. (1967). Biochim.Biophys.Acta. 133, 195.

Shotton, D.M. & Hartley, B.S. (1970). Nature 225, 802.

Shotton, D.M. & Watson, H.C. (1970). Nature 225, 811.

Sia, C. & Horecker, B.L. (1968a). Arch.Biochem.Biophys. 123, 186.

Sia, C. & Horecker, B.L. (1968b). Biochem.Biophys.Res.Commun. 31, 731.

Sibley, J.A. & Lehninger, A.L. (1949). J.Biol.Chem. 177, 859.

Sigler, P.B. & Blow, D.M. (1965). J.Mol.Biol. 14, 640.

Signor, A., Biondi, L., Tamburro, A.M. & Bordignon, E. (1969). Eur.J.Biochem. 7, 328.

Simmons, R.M. (1965). Ph.D. Thesis presented to London University.

Sine, H.E. & Hass, L.F. (1967). J.Am.Chem.Soc. 89, 1749.

Singh, A.K. & Ramaseshan, S. (1966). Acta Cryst. 21, 279.

Sørensen, S.P.L. & Sorensen, M. (1933). Biochem.Z. 258, 16.

Speck, J.C., Rowley, P.T., & Horecker, B.L. (1963). J.Am.Chem. Soc. 85, 1012.

Spolter, P.D., Adelman, R.C. & Weinhouse, S. (1965). J.Biol.Chem. 240, 1327.

Steinrauf, L.K. (1963). Acta Cryst. 16, 317.

Stellwagen, E. & Schachman, H.K. (1962). Biochemistry I, 1056.

Susi, H., Timasheff, S.N. & Stevens, L. (1968). J.Biol.Chem. 242, 5460.

Susi, H., Zell, T. & Timasheff, S.N. (1959). Arch.Biochem.Biophys. 85, 437.

Susor, W.A., Kochman, M. & Rutter, W.J. (1969). Science. 165, 1260.

Swenson, A.D. & Boyer, P.D. J.Am.Chem.Soc. 79, 2174. (1957).

Szabolsci, E., & Biszku, E. (1961). Biochim.Biophys. Acta. 48, 335.

Szajani, B., Sajgo, M., Bisku, E., Friedrich, P. & Szabolcsi, G. (1970). Eur. J.Biochem. 15, 171.

Tanford, C., Bunville, L.G. & Nozaki, Y. (1959). J.Am.Chem.Soc. 81, 4032.

Tanford, C. & Nozaki, Y. (1959). J.Biol.Chem. 234, 2874.

Tanford, C. & Taggart, V.G. (1961). J.Am.Chem.Soc. 83, 1634.

Tanford, C., De, P.K. & Taggart, V.G. (1960). J.Am.Chem.Soc. 82, 6028.

Taylor, J.F. (1955). Methods in Enzymology. Vol.I. Eds. Colowick, S.P. & Kaplan, N.O. Academic Press, New York.

Taylor, J.F. (1957). Biochemical Preps. 5, 12.

Taylor, J.F., Green, A.A. & Cori, G.T. (1948). J.Biol.Chem, 173, 591.

Timasheff, S.N., Mescanti, L., Basch, J.J. & Townend, R. (1966). J.Biol.Chem. 241, 2496.

Timasheff, S.N. & Townend, R. (1961a). J.Am.Chem.Soc. 83, 464.

Timasheff, S.N. & Townend, R. (1961b). J.Am.Chem.Soc. 83, 470.

Townend, R. (1965). Arch.Biochem.Biophys. 109, 1

Townend, R., Herskovits, T.T., Timasheff, S.N. & Gorbunoff, M.J. (1969). Arch.Biochem.Biophys. 129, 567.

Townend, R., Kumosinski, T.F. & Timasheff, S.N. (1968). J.Biol.Chem. 242, 4538.

Townend, R., Herskovits, T.T., Swaisgood, H.E. & Timasheff, S.N. (1964). J.Biol.Chem. 239, 4196.

Townend, R., Kiddy, C.A. & Timasheff, S.N. (1961). J.Am.Chem.Soc. 83, 1419.

Townend, R. & Timasheff, S.N. (1957). J.Am.Chem.Soc. 79, 3613.

Udenfreund, S. & Velick, S.F. (1951). J.Biol.Chem. 190, 733.

Velick, S.F. & Ronzoni, E. (1948). J.Biol.Chem. 173, 627.

Vogel, A.I. (1962). A Textbook of Quantitative Inorganic Analysis Longmans, London.

Warburg, O. & Christian, W. (1943). Biochem.Z. 314, 149.

Watson, H.C., Shotton, D.M., Cox, J.M. & Muirhead, H. Nature 225, 806.

Webb, J.L. (1963). Enzyme and Metabolic Inhibitors, Vol. I p.839. New York: Academic Press.

Westhead, E.W., Butler, L. & Boyer, P.D. (1963). Biochemistry 2, 927.

Wilson, A.J.C. (1942). (Ed.) Structure Reports 9, 207.

Winstead, J.A. & Wold, F. (1964). J.Biol.Chem. 239, 4212.

Witz, J., Timasheff, S.N. & Luzzati, V. (1964). J.Am.Chem.Soc. 86, 168.

Wolf, H.P. & Leuthardt, F. (1957). Helv. Chim. Acta. 40, 237.

Wyckoff, R.W.G. (1965). Crystal Structures, Vol III, 2nd Edn p.686, Interscience.

ABSTRACT OF THESIS

Name of Candidate Lindsay Sawyer

Address 73, Promenade, James Street, Edinburgh, EH15 2DX

Degree Ph.D. Date June 1971

Title of Thesis Structural Studies of Two Proteins

This work can be divided into two parts: in the first, the crystallisation and preliminary X-ray data for rabbit muscle aldolase (E.C. 4.1.2.13) is examined, in the second, the search for suitable high resolution derivatives for bovine β-lactoglobulin is studied with a view to the correlation of solution and X-ray studies of the binding of gold, platinum and iridium complexes to the protein.

Rabbit muscle aldolase from the back and leg muscles of freshly killed rabbits can be prepared by ammonium sulphate fractionation as a microcrystalline suspension. At least two forms of crystal exist: at pH values below 7.0 the crystals are hexagonal bipyramids, above 7.0 hexagonal plates can be formed. Crystals of a suitable size for X-ray work were grown from 2.0M phosphate, pH 6.0 in the form of hexagonal bipyramids. Precession photographs along and perpendicular to the sixfold axis allowed the cell dimensions and space group to be determined as a=163.5 Å, c=335.0 Å and $P6_122$. There were 18 molecules in the unit cell or 1.5 in the asymmetric unit implying that the molecule is a tetramer rather than a trimer. Although crystals of the high pH form are obtained they are not of sufficient quality to allow X-ray study.

With β-lactoglobulin, the binding of tetracyanoaurate (III) is examined in solution by ultraviolet spectroscopy and polarimetry and also by X-rays in the solid phase. In solution it is found that one mole of heavy atom complex per subunit binds specifically and reversibly, $\text{Protein} - SH + Au(CN)_4^- \rightleftharpoons \text{Protein} - S - Au(CN)_3^- + HCN$, to the free cysteine causing a marked change in specific rotation, to a stable conformation midway between the R and S states. The effect of concentrated salt is to inhibit this change and it is proposed that lattice X crystals correspond to the N state, lattices Y and Z to an intermediate one and that the R state should be obtained by crystallising the β-lactoglobulin-gold complex. X-ray studies on lattice Y show that the binding occurs mainly at the free sulphydryl group and also at the $HgI_4^=$ site but in lattice Z, no sulphydryl binding is detected. The derivatives are

Use other side if necessary.

markedly non-isomorphous.

With the isostructural tetracyanoplatinite (II), no binding to the free sulphydryl group is detected either in solution or by X-ray crystallography. The electron density map shows several minor sites.

Further studies with hexacyanoplatinate (IV) and the chlorocomplexes of platinum and iridium show that the iridium complexes look promising as high resolution derivatives although no satisfactory correlation is obtained between solution and crystal studies. The chlorocomplexes of platinum do show evidence of binding in solution but they appear to react to ratios greater than one to one. No satisfactory X-ray data are obtained in lattice Z with these derivatives.

In conclusion, a strategy for the crystallisation of proteins with a view to the requirements of the crystallographer is proposed.